TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA ĐIỆN – ĐIỆN TỬ**



**LÊ PHƯƠNG ANH**

**ĐIỀU KHIỂN VÀ GIÁM SÁT THIẾT BỊ BAY KHÔNG NGƯỜI LÁI DẠNG QUADCOPTER**

**ĐỒ ÁN CHUYÊN NGÀNH**

**KỸ THUẬT ĐIỀU KHIỂN VÀ TỰ ĐỘNG HÓA**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024**

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA ĐIỆN – ĐIỆN TỬ**



**LÊ PHƯƠNG ANH**

**ĐIỀU KHIỂN VÀ GIÁM SÁT THIẾT BỊ BAY KHÔNG NGƯỜI LÁI DẠNG QUADCOPTER**

**ĐỒ ÁN TỔNG HỢP**

**KỸ THUẬT ĐIỀU KHIỂN VÀ TỰ ĐỘNG HÓA**

Người hướng dẫn

**TS. Vũ Trí Viễn**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024**

**LỜI CẢM ƠN**

Tôi xin chân thành cảm ơn TS. Vũ Trí Viễn đã dành thời gian để hướng dẫn tôi hoàn thành đồ án một cách tận tình và nhiệt tình nhất. Bên cạnh đó, tôi cũng cảm ơn một số thầy cô trong khoa Điện – Điện tử, các anh chị đi trước và bạn bè đã chỉ dẫn tôi trong quá trình hoàn thành đồ án này. Đây là đồ án chuyên ngành nên có nhiều sai phạm và thiếu sót nên tôi hy vọng nhận được sự góp ý từ thầy cô trong khoa để có thể rút kinh nghiệm cho các đồ án sắp tới.

Tôi xin gửi lời cảm ơn sâu sắc đến thầy cô, anh chị và bạn bè giúp đỡ tôi trong đồ án này!

|  |  |
| --- | --- |
|  | *TP. Hồ Chí Minh, ngày 10 tháng 1 năm 2024* |
|  | *Tác giả* |
|  |  |
|  | Lê Phương Anh |

Công trình được hoàn thành tại Trường Đại học Tôn Đức Thắng

Cán bộ hướng dẫn khoa học: Tiến sĩ Vũ Trí Viễn

........................................................................................

*(Ghi rõ học hàm, học vị, họ tên và chữ ký)*

Đồ án chuyên ngành được bảo vệ tại **Hội đồng đánh giá Đồ án chuyên ngành của Trường Đại học Tôn Đức Thắng** vào ngày… /…/……

Xác nhận của Chủ tịch Hội đồng đánh giá Đồ án chuyên ngành và Trưởng khoa quản lý chuyên ngành sau khi nhận Đồ án chuyên ngành đã được sửa chữa (nếu có).

**CHỦ TỊCH HỘI ĐỒNG TRƯỞNG KHOA**

**…………………………. ………………………………**

**CÔNG TRÌNH ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi và được sự hướng dẫn khoa học của **TS. Vũ Trí Viễn**. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong Đồ án chuyên ngành còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung Đồ án chuyên ngành của mình.** Trường Đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

|  |  |
| --- | --- |
|  | *TP. Hồ Chí Minh, ngày 10 tháng 1 năm 2024* |
|  | *Tác giả* |
|  |  |
|  | Lê Phương Anh |

|  |  |
| --- | --- |
| TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG  **KHOA ĐIỆN –ĐIỆN TỬ**  ------------------- | CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM  Độc lập – Tự do – Hạnh phúc  ---------------------- |

**LỊCH TRÌNH LÀM ĐỒ ÁN CHUYÊN NGÀNH**

Họ tên sinh viên: Lê Phương Anh

Lớp:19040303 MSSV: 41900688

Tên đề tài: Phân tích phổ âm thanh STFT

|  |  |  |  |
| --- | --- | --- | --- |
| **Tuần/Ngày** | **Khối lượng** | | **GVHD ký** |
| **Đã thực hiện** | **Tiếp tục thực hiện** |
| Tuần 1 (29/08) | Nhận đề tài Đồ án chuyên ngành | Tìm hiểu đề tài và đọc tài liệu tham khảo liên quan |  |
| Tuần 2 (05/09) | Tìm hiểu đề tài và đọc tài liệu tham khảo liên quan | Tìm hiểu Mic array, giải thuật STFT |  |
| Tuần 3 (12/09) | Tìm hiểu Mic array, giải thuật STFT | Tìm hiểu hướng âm thanh và áp dụng để thu âm thanh thông qua phần cứng |  |
| Tuần 4 (19/09) | Tìm hiểu hướng âm thanh và áp dụng để thu âm thanh thông qua phần cứng | Tìm hiểu node red, node.js |  |
| Tuần 5 (26/09) | Tìm hiểu node red, node.js | Bắt đầu triển khai chương trình |  |
| Tuần 6 (03/10) | Tìm được các file mẫu | Xây dựng khối model |  |
| Tuần 7 (10/10) | Xây dựng được khối model | Xây dựng khối train data |  |
| Kiểm tra giữa kỳ | Đánh giá khối lượng hoàn thành……..%  được tiếp tục/không tiếp tục thực hiện ĐACN | | |
| Tuần 8 (24/10) | Xây dựng khối train data | Xây dựng khối predict |  |
| Tuần 9 (31/10) | Xây dựng được khối predict lấy dữ liệu từ file có sẵn và micro | Chạy và sửa lỗi chương trình |  |
| Tuần 10 (07/11) | Chạy và sửa lỗi chương trình | Viết báo cáo đồ án |  |
| Tuần 11 (14/11) | Hoàn thành ½ báo cáo | Viết tiếp báo cáo và làm poster |  |
| Tuần 12 (21/11) | Hoàn thành báo cáo và poster | Hoàn thanh đồ án |  |
| Nộp Đồ án hệ thống nhúng | Đã hoàn thành……..% Đồ án chuyên ngành  được bảo vệ/không được bảo vệ ĐACN | | |

**PHÂN TÍCH PHỔ ÂM THANH STFT**

**TÓM TẮT**

Nhiệm vụ:

- Tìm hiểu micro array

- Tìm hiểu giải thuật stft

- Thiết kế mạch thu và nhận dạng hướng âm thanh

- Viết báo cáo và làm giao diện laptop

MỤC LỤC

DANH MỤC HÌNH VẼ 3

CHƯƠNG 1. TỔNG QUAN VỀ ĐỀ TÀI NGHIÊN CỨU 5

1.1 Giới thiệu về đề tài 5

1.2 Mục đích nghiên cứu 5

1.3 Đối tượng nghiên cứu 5

1.4 Phạm vi nghiên cứu 6

1.5 Dự kiến kết quả 6

CHƯƠNG 2. CƠ SỞ LÝ THUYẾT 7

2.1 Sơ lược về phổ âm thanh 7

2.2 Biến đổi Fourier nhanh (FFT) 8

2.3 Tìm hiểu giải thuật stft 9

2.4 Tìm hiểu về micro array 12

2.5 Cấu trúc mạng nơ-ron tích chập (Convolutional Neural Network – CNN) 14

CHƯƠNG 3. THIẾT KẾ VÀ THI CÔNG 16

3.1 Sơ đồ khối của hệ thống phân loại âm thanh phổi 16

3.1.1 Khối âm thanh phổi 16

3.1.2 Khối xử lý tín hiệu 17

3.1.3 Khối trích xuất đặc trưng 18

3.1.4 Khối phân loại âm thanh 18

3.2 Sơ đồ khối của hướng âm thanh 18

3.2.1 Bộ lọc hệ thống AEC (Acoustic Echo Cancellation) 18

3.2.2 DOA (Direction of Arrival) 19

3.2.3 Beamforming 20

3.2.4 NS (Noise Suppression) 21

3.2.5 AGC (Automatic Gain Control) 21

3.3 Thiết kế giao diện 22

CHƯƠNG 4. GIẢI THUẬT VÀ ĐIỀU KHIỂN 23

4.1 Hoạt động của hệ thống 23

4.1.1 Phân loại âm thanh phổi 23

4.1.2 Hướng âm thanh 23

4.2 Lưu đồ giải thuật 24

4.2.1 Phân loại âm thanh phổi 24

4.2.2 Hướng âm thanh 24

CHƯƠNG 5. THỰC NGHIỆM 25

5.1 Tiến trình thực nghiệm 25

5.1.1 Phân loại âm thanh phổi 25

5.1.2 Hướng âm thanh 27

5.2 Kết quả thực nghiệm 28

5.2.1 Kết quả đào tạo 28

5.2.2 Kết quả hướng âm thanh 30

5.2.3 Kết quả dự đoán 30

5.3 Kết luận thực nghiệm 35

CHƯƠNG 6. KẾT LUẬN 36

6.1 Ưu điểm 36

6.2 Nhược điểm 36

6.3 Hướng phát triển 36

TÀI LIỆU THAM KHẢO 37

PHỤ LỤC A CHƯƠNG TRÌNH HƯỚNG ÂM THANH 1

# DANH MỤC HÌNH VẼ

Hình 1. Phổ âm thanh 7

Hình 2. Chuyển đổi tín hiệu tương tự sang kỹ thuật số 7

Hình 3. Chuyển đổi từ miên thời gian sang miền tần số dùng FFT 8

Hình 4. Cấu trúc thuật toán FFT 9

Hình 5. Mô tả tín hiệu tiếng nói 9

Hình 6. Quang phổ 10

Hình 7. Chuyển dổi STFT từ biên độ sang tần số 11

Hình 8. Cửa sổ Kaiser 12

Hình 9. Micro array 12

Hình 10. Hướng âm thanh 13

Hình 11. Độ nhạy micro 13

Hình 12. Độ lệch pha giữa hai sóng âm 14

Hình 13. Cấu trúc của CNN 15

Hình 14. Sơ đồ khối của hệ thống 16

Hình 15. Hình Phân tích phổi âm thanh giữa phổi bình thường, Crackle và Wheeze. (a) Wavelet ban đầu. (b) Phổ STFT của âm thanh phổi. (c) Ma trận phân tích sóng của âm thanh phổi. 17

Hình 16. Sơ đồ khối hướng âm thanh 18

Hình 17. Ước tính DOA 20

Hình 18. Hai mẫu chùm tia 20

Hình 19. Tín hiệu vào (a) và tín hiệu sau khi lọc nhiễu (b) 21

Hình 20. Sơ đồ của một AGC được sử dụng trong mạng điện thoại analog 21

Hình 21. Giải thuật trên node red 22

Hình 22. Giao diện dashboard của node red 22

Hình 23. Lưu đồ giải thuật của hệ thống phân loại âm thanh phổi 24

Hình 24. Lưu đồ giải thuật hướng âm thanh 24

Hình 25. Setup Parameter trong block Build Model to Train AI 25

Hình 26. Setup Parameter trong block Train Data 26

Hình 27. Block Build Model to Train AI và block Train Data 26

Hình 28. Các cách đưa dữ liệu vào 26

Hình 29. Mạch thu âm thanh thực tế 27

Hình 30. Chi tiết mạch thu âm thanh 27

Hình 31. Model thu được sau khi xây dựng 28

Hình 32. Thông số cuối cùng của 100 epoch 28

Hình 33. Thông số cuối cùng của 500 epoch 29

Hình 34. Thông số cuối cùng của 1000 epoch 29

Hình 35. Âm thanh thu được sau khi được xử lý thông qua mạch 30

Hình 36. Tín hiệu đầu vào micro (người khỏe mạnh) cho ra kết quả không ho 30

Hình 37. Tín hiệu đầu ra của micro (không ho) 30

Hình 38. Tín hiệu đầu vào file normal (2) cho ra kết quả không ho 31

Hình 39. Tín hiệu đầu ra của file normal 31

Hình 40. Tín hiệu đầu vào file crack (2) cho ra kết quả có ho 31

Hình 41. Tín hiệu đầu ra của file crack 31

Hình 42. Tín hiệu đầu vào micro (người khỏe mạnh) cho ra kết quả không ho 32

Hình 43. Tín hiệu đầu ra của micro (không ho) 32

Hình 44. Tín hiệu đầu vào file normal (5) cho ra kết quả không ho 32

Hình 45. Tín hiệu đầu ra của file normal 32

Hình 46. Tín hiệu đầu vào file crack (5) cho ra kết quả có ho 33

Hình 47. Tín hiệu đầu ra của file crack 33

Hình 48. Tín hiệu đầu vào micro (người khỏe mạnh) cho ra kết quả không ho 33

Hình 49. Tín hiệu đầu ra của micro (không ho) 33

Hình 50. Tín hiệu đầu vào file normal (3) cho ra kết quả không ho 34

Hình 51. Tín hiệu đầu ra của file normal 34

Hình 52. Tín hiệu đầu vào file crack (3) cho ra kết quả có ho 34

Hình 53. Tín hiệu đầu ra của file crack 34

# TỔNG QUAN VỀ ĐỀ TÀI NGHIÊN CỨU

## Giới thiệu về đề tài

Trong thời buổi hiện nay, nhận dạng âm thanh cũng là một trong số ứng dụng dùng Iot, AI, nhờ vào đó mà máy móc có thể giúp con người xử lý những dạng âm thanh khác nhau để phân loại theo mục đích sử dụng chúng. Muốn nhận dạng được những luồng âm thanh đó thì chúng ta cần phải phân tích được phổ âm thanh của chúng.

Một trong số những lĩnh vực được áp dụng AI phổ biến đó chính là lĩnh vực y học, dùng máy móc để phát hiện ra nhiều căn bệnh phức tạp một cách nhanh chóng. Phổ biến nhất trong thời buổi này đó chính là những căn bệnh liên quan đến hô hấp vì nạn ô nhiễm và cũng như là dịch bệnh hoành hành. Dựa vào những lý do đó mà chúng ta cần phát triển thêm ở mảng này để có thể dễ dàng nhận ra giữa một người có phổi bình thường và người có phổi bất thường để kịp thời cứu chữa. Và đó cũng chính là đề tài của em “Phân tích phổ âm thanh STFT” dùng để phân loại giữa âm thanh phổi bình thường và phổi bất thường.

## **Mục đích nghiên cứu**

Ứng dụng các kiến thức đã học (xử lý số tín hiệu,…) để phân tích phổ âm thanh STFT.

Hiểu rõ cách thức hoạt động của thuật toán Deep learning, cụ thể là cấu trúc mạng CNN.

Phân loại được giữa âm thanh phổi người khỏe mạnh với âm thanh phổi người bệnh.

## **Đối tượng nghiên cứu**

* Âm thanh phổi (bình thường và bất thường)
* Phổ âm thanh
* Giải thuật stft
* Micro array

## **Phạm vi nghiên cứu**

Dùng để áp dụng trong các ứng dụng trong lĩnh vực y học, phân loại các hướng âm thanh.

## **Dự kiến kết quả**

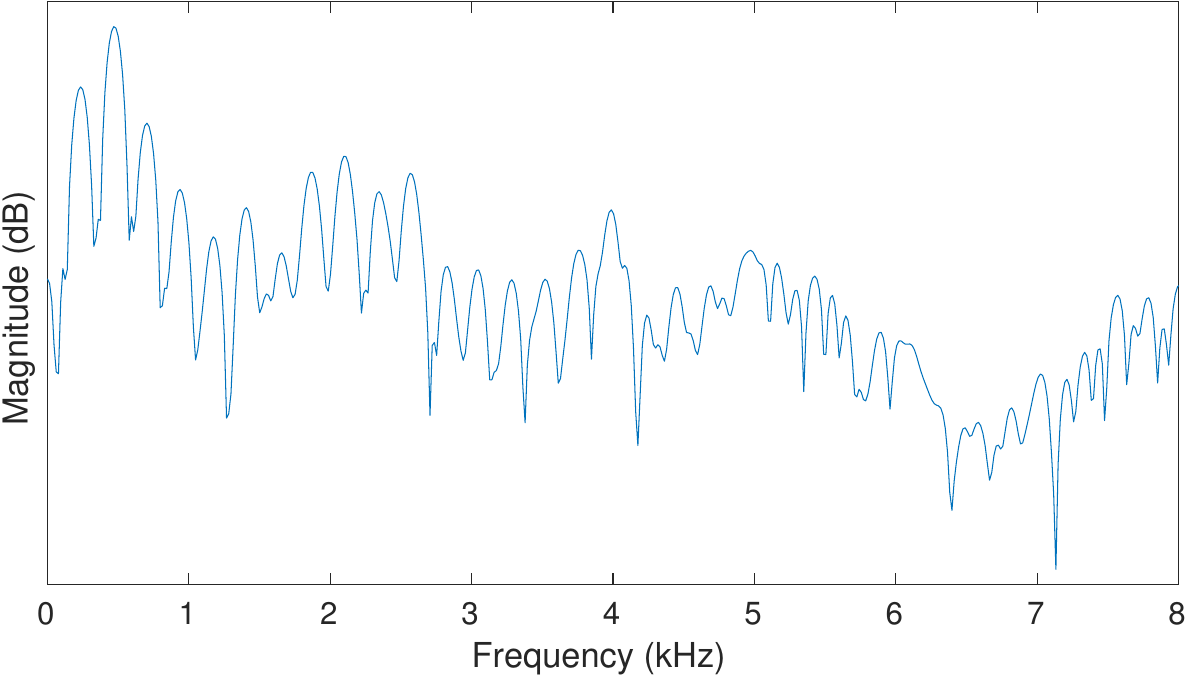
* Cho ra được giao diện của laptop
* Thiết kế được mạch thu và nhận dạng hướng âm thanh
* Phân loại được giữa âm thanh phổi của người khỏe mạnh và âm thanh phổi bất thường

# CƠ SỞ LÝ THUYẾT

## Sơ lược về phổ âm thanh

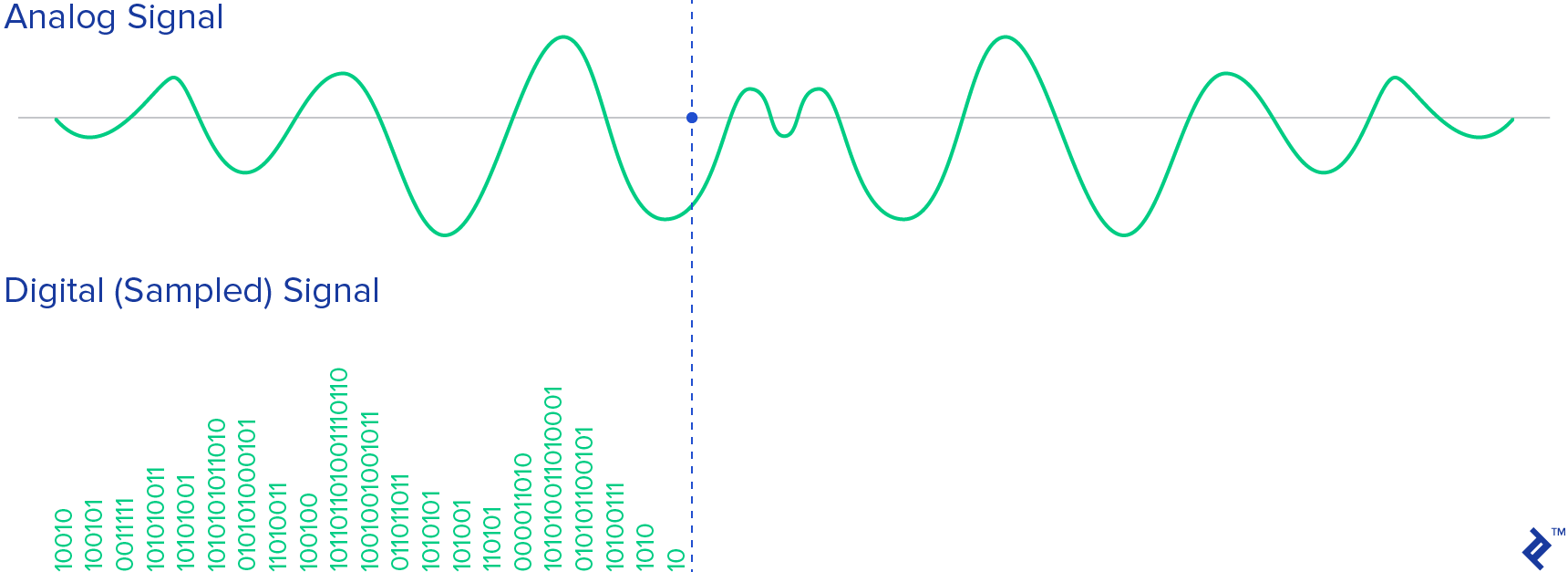
Sự phân bố âm phổ là những âm thanh mà ta nghe được từ bất kỳ vật thể nào phát ra và nó được tạo thành từ hỗn hợp nhiều sóng sine với nhau.

Phổ âm thanh hiển thị các tần số khác nhau có trong âm thanh, nó cũng được xem là đại diện cho âm thanh – là một mẫu ngắn của âm thanh – về số lượng sóng sine ở từng tần số riêng lẻ. Nó thường được hiển thị thành một biểu đồ của tần số và đo bằng nhiều rung động tại mỗi giây.

Lấy ví dụ về phổ âm thanh

1. Phổ âm thanh

Ngày nay, phổ âm thanh thường đo bằng cách sử dụng:

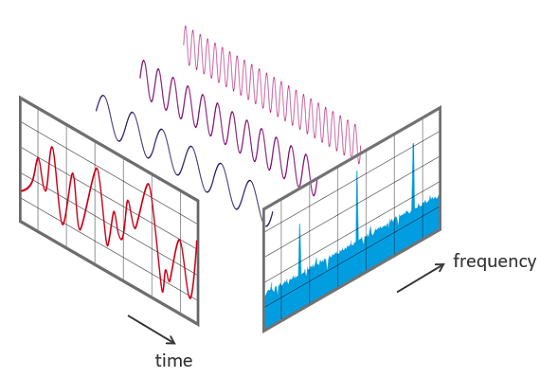
* Một Micro đo áp suất âm thanh qua một khoảng thời gian nhất định.
* Một công cụ chuyển đổi tín hiệu tương tự sang kỹ thuật số được thực hiện nhiều chuyển đổi trên các phần rất nhỏ của tín hiệu, quá trình này còn được gọi là lấy mẫu.

1. Chuyển đổi tín hiệu tương tự sang kỹ thuật số

* Một máy tính thực hiện tính toán dựa trên những con số đó.

## Biến đổi Fourier nhanh (FFT)

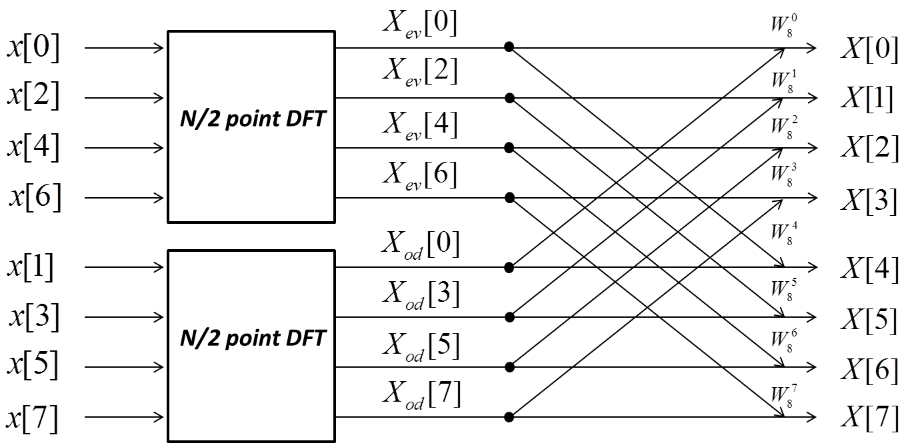
Khi xử lý bằng FFT (Fast Fourier Transform), ngoài việc giảm bớt được thời gian tính toán thì nó còn làm giảm đáng kể các lỗi hệ thống liên quan đến các tính toán. Thời gian tính toán và lỗi hệ thống thường bị giảm bởi hệ số . Trong đó, N đại diện cho số lượng mẫu dữ liệu trong chuỗi thời gian.

Biến đổi Fourier nhanh là một kỹ thuật kết hợp tuần tự các tổng dữ liệu có trọng số lớn hơn dần dần do đó tạo ra các hệ số DFT cần thiết. Thuật toán FFT được sử dụng để tích chập nhanh (tuyến tính, bộ lọc biến đổi bất biến). FFT có rất nhiều ứng dụng và được sử dụng rộng rãi trong xử lý âm thanh, radar, sonar,… Trong tất cả các ứng dụng này, tín hiệu miền thời gian nhờ vào FFT có thể chuyển đổi và được biểu diễn ở miền tần số của tín hiệu.

1. Chuyển đổi từ miên thời gian sang miền tần số dùng FFT

FFT của được xác định bởi công thức:

Trong đó, h(n) là cửa sổ Hamming

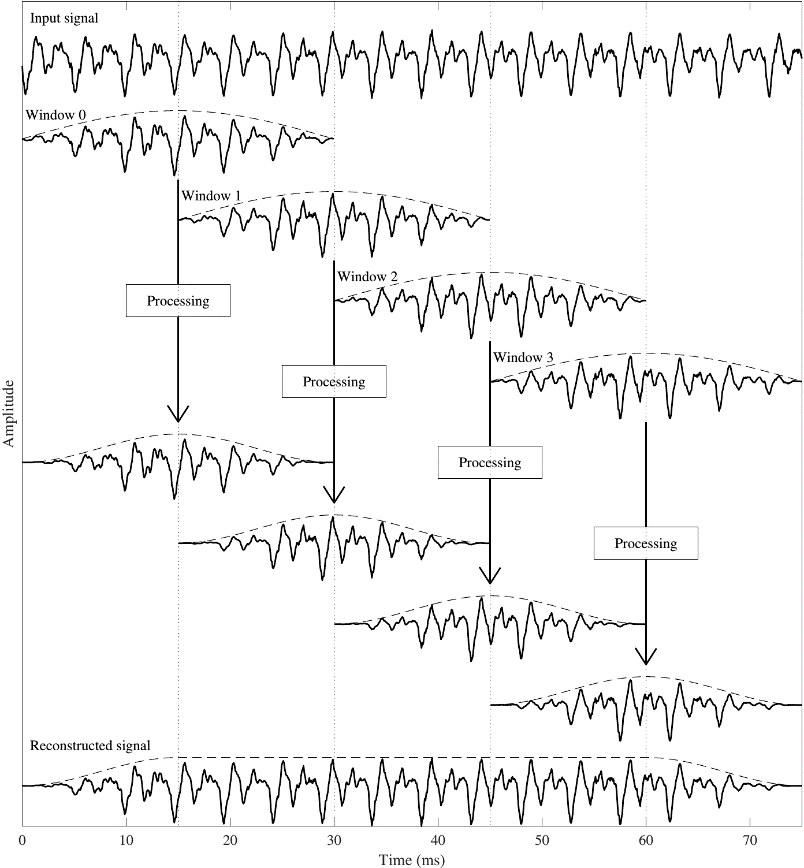


1. Cấu trúc thuật toán FFT

## Tìm hiểu giải thuật stft

Còn được gọi là biến đổi fourier thời gian ngắn là sự phân chia một chuỗi thời gian thành các khối chồng nhau có chiều dài bằng nhau và áp dụng biến đổi Fourier nhanh (FFT) cho mỗi khối một cách tuần tự.

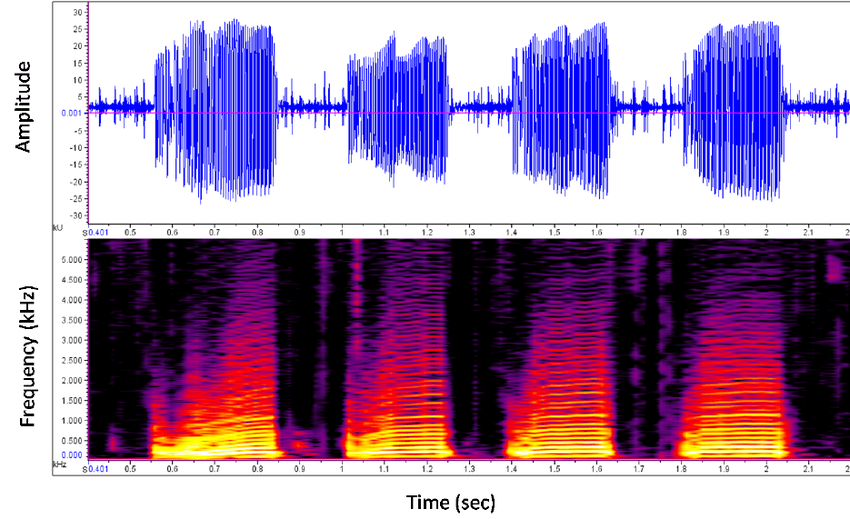
Do tín hiệu tiếng nói là tín hiệu không đứng yên, nên không thể áp dụng phép phân tích Fourier thông thường. Song, nếu chúng ta chia tín hiệu tiếng nói ra thành từng đoạn đủ nhỏ theo thời gian, thì tín hiệu tiếng nói trong mỗi đoạn có thể xem là tín hiệu đứng yên, và do đó có thể lấy biến đổi Fourier trên từng đoạn tín hiệu này. Ðây là nguyên lý có phép biến đổi Fourier thời gian ngắn, còn gọi là biến đổi Fourier cửa sổ hóa.



1. Mô tả tín hiệu tiếng nói

STFT là một trong những công cụ được sử dụng thường xuyên nhất trong phân tích và xử lý giọng nói. Nó mô tả sự phát triển của các thành phần tần số theo thời gian. Giống như quang phổ, một trong những lợi ích của STFT là các tham số của nó có cách giải thích vật lý và trực quan.

Một điểm tương đồng nữa với quang phổ là đầu ra của STFT có giá trị phức tập, mặc dù trong đó đầu ra phổ là một vecto, đầu ra STFT là một ma trận. Kết quả là, chúng ta không thể trực tiếp nhìn thấy đầu ra có giá trị phức tạp. Thay vào đó, STFT thường được hiển thị bằng cách sử dụng quang phổ log. Quang phổ log 2 chiều sau đó được hiển thị bằng bản đồ nhiệt được gọi là quang phổ, có công thức như sau:

20log10(*X*(*h*,*k*))

1. Quang phổ

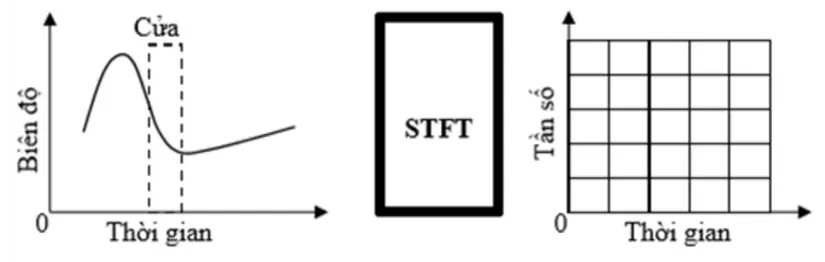
Đầu tiên tín hiệu được nhân với một hàm cửa sổ và sau đó thực hiện biến đổi Fourier, kết quả sẽ cho một biến đổi 2 chiều STFT:

Trong biến đổi Fourier thời gian ngắn (STFT) các hàm sử dụng trong mở rộng thu được bằng cách làm trễ và điều chỉnh hàm cửa sổ cơ sở :

Từ đó có được một dạng mở rộng:

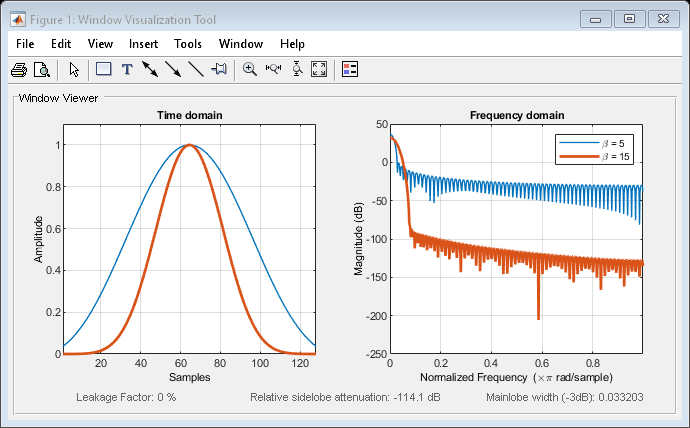
Ðể thấy rõ STFT cũng định vị trong miền tần số, ta có thể áp dụng định lý Parserval:

Với W(') và F(') lần lượt là phổ của cửa sổ w(t) và của tín hiệu f(t). Trong biểu thức cuối cùng, W( '-) có tác dụng như một lọc dãy thông tập trung quanh tần số đang phân tích và có băng thông bằng với băng thông của w(t), làm giới hạn phổ của tín hiệu F(') xung quanh . STFT có tính định vị theo tần số. Tính định vị này (còn gọi là độ phân giải tần số) càng tốt nếu băng thông của cửa sổ phân tích càng hẹp.

1. Chuyển dổi STFT từ biên độ sang tần số

Hàm cửa sổ thường dùng trong STFT là cửa sổ Kaiser, hàm này được định nghĩa từ hàm Bessel bậc 0:

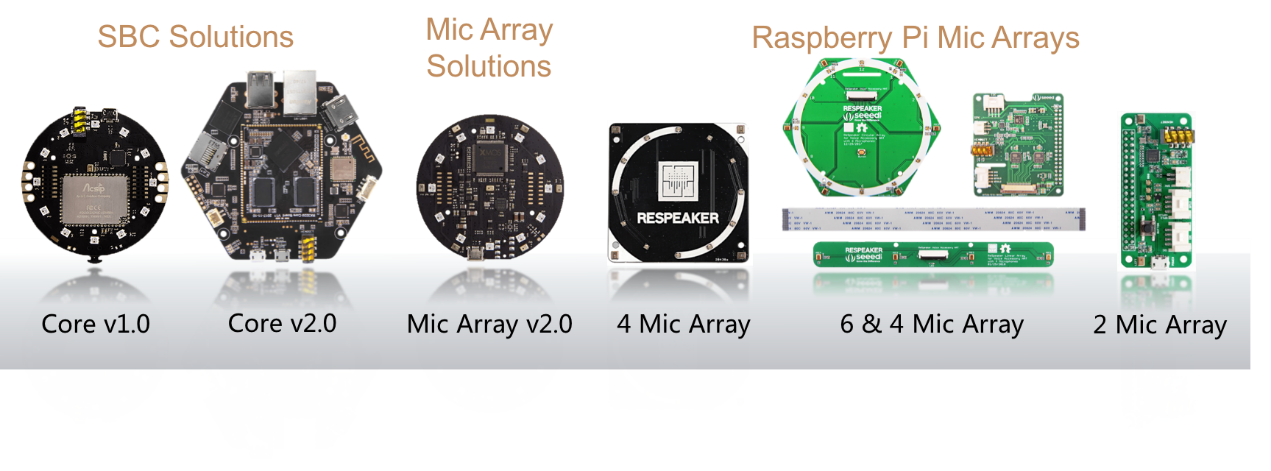
Hàm Kaiser có thể thay đổi linh hoạt nhờ vào thông số hình dạng (shape parameter) . Với các giá trị khác nhau, cửa sổ Kaiser sẽ có hình dạng khác nhau.

1. Cửa sổ Kaiser

## Tìm hiểu về micro array

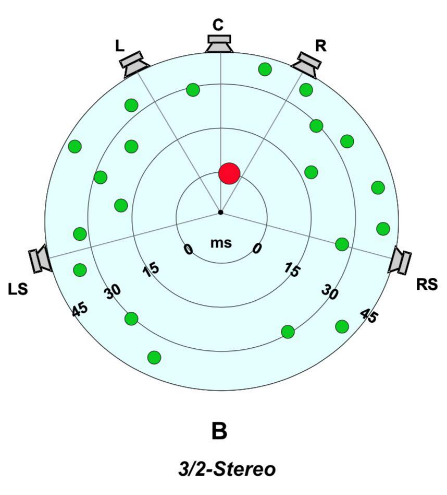
Micro array (mảng micro) là một thiết bị micro hoạt động giống như một thiết bị micro thông thường, nhưng thay vào đó chỉ có môt micro để ghi lại đầu vào âm thanh, nó có nhiều micro (2 hoặc nhiều hơn) để ghi lại âm thanh.

Các micro trong mảng hoạt động đồng thời cùng nhau để ghi lại âm thanh. Đặc điểm quan trọng nất phải có trong các thiết bị mảng micro là sự kết hợp của nhiều micro với nhau. Tất cả các micro trong mảng phải giống nhau để tín hiệu nhận được giữa các micro có sự tương đồng.

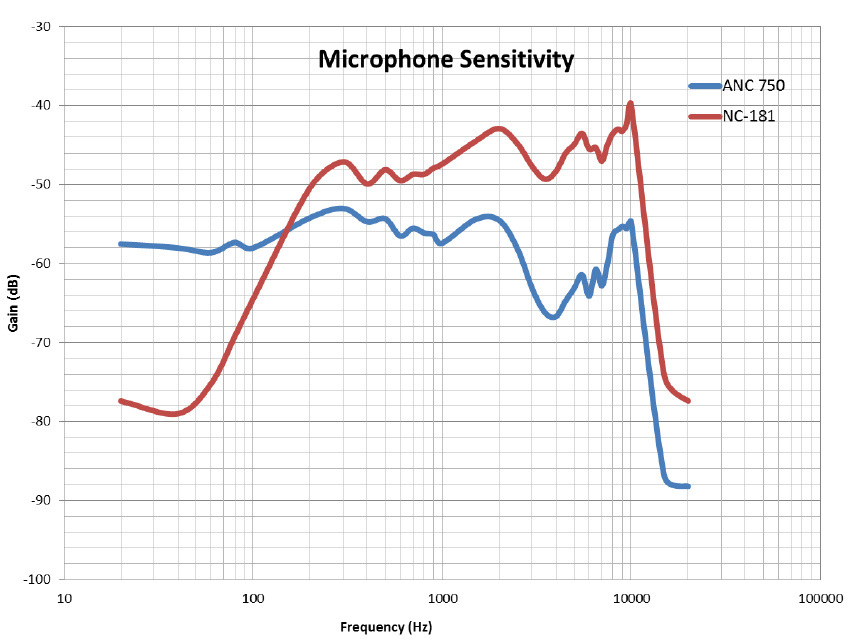


1. Micro array

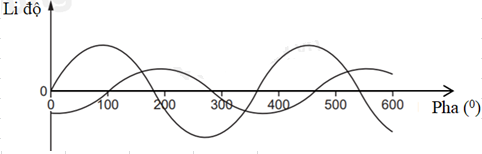
Có ba yếu tố cần xem xét để kết hợp micro trong mảng micro là hướng âm thanh, độ nhạy và pha.

* Hướng âm thanh là hướng mà nó thu được âm thanh. Một số micro được tạo ra chỉ để nhận âm thanh từ một hướng, một chiều micro. Số khác được chế tạo ra để chúng có thể thu âm thanh từ mọi hướng còn được gọi là micro đa hướng. Khi xây dựng mảng micro, tất cả micro phải có cùng hướng.

1. Hướng âm thanh

* Độ nhạy là một yếu tố khác phải phù hợp với mảng micro. Độ nhạy phải được kết hợp chặt chẽ trong các thiết bị mảng micro, nếu không micro này sẽ to hơn micro kia, tạo ra các bản ghi mất cân bằng. Chênh lệch độ nhạy tối đa cho phép trong mảng micro là ±1.5dB.

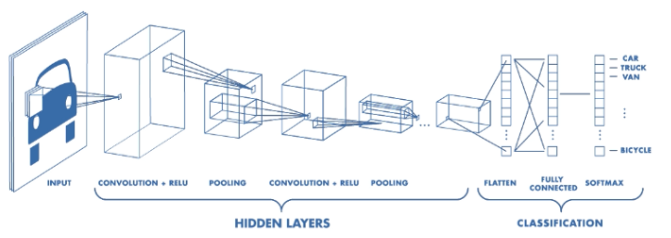
1. Độ nhạy micro

* Độ lệch pha là yếu tố quan trọng cuối cùng của một mảng micro phải được phù hợp. Pha là độ tham chiếu dòng khi micro bắt đầu ghi, nó xác định khi nào mà tất cả các micro cùng một mảng bắt đầu và dừng ghi. Nếu các micro có các pha khác nhau thì chúng sẽ ghi lại tín hiệu ở các thời điểm khác nhau. Điều này sẽ khiến việc ghi âm không đồng bộ, đó là việc không mong muốn. Điều mong muốn là khi ghi lại các tín hiệu đầu vào đồng thời thì sẽ không có độ trễ giữa các tín hiệu. Cũng giống như độ nhạy, độ lệch pha giữa các micro cho phép trong khoảng ±1.5 độ để đảm bảo rằng các tín hiệu được ghi lại đồng thời.
* 

1. Độ lệch pha giữa hai sóng âm

## Cấu trúc mạng nơ-ron tích chập (Convolutional Neural Network – CNN)

CNN là một tập con của Machine Learning và là trung tâm của các thuật toán Deep Learning. Nó bao gồm các lớp: lớp tích chập (Convolution), lớp lấy mẫu (Pooling), lớp kết nối đầy đủ (Fully Connected) và lớp phi tuyến tính ReLU (Rectified Linear Unit)

* Lớp tích chập là khối xây dựng cốt lõi của CNN. Lớp này yêu cầu một vài thành phần như dữ liệu đầu vào, bộ lọc và bản đồ tính năng. Khi đưa dữ liệu vào thì nhờ vào các tính năng hoặc bộ lọc sẽ di chuyển qua các trường tiếp nhận của hình ảnh, quá trình này được gọi là một kết cấu. Đầu ra cuối cùng từ loạt sản phẩm chấm từ đầu vào và bộ lọc được gọi là tính năng tích chập.
* Lớp lấy mẫu là lấy mẫu các lớp, còn được gọi là downsampling, được tiến hành để giảm chiều hướng, giảm số lượng tham số cho đầu vào. Tương tự như lớp tính chập, thao tác gộp quét một bộ lọc cho toàn bộ đầu vào, nhưng sự khác biệt là bộ lọc này không có bất kỳ trọng lượng nào. Có hai loại lấy mẫu chính: Max pooling (lấy mẫu tối đa) và Average pooling (lấy mẫu trung bình).
* Lớp kết nối đầy đủ: lớp này thực hiện nhiệm vụ phân loại dựa trên các tính năng được trích xuất thông qua các lớp trước đó và các bộ lọc khác nhau của chúng. Trong khi các lớp tích chập và lấy mẫu có xu hướng sử dụng các hàm ReLU thì lớp kết nối đầy đủ thường tận dụng chức năng kích hoạt softmax để phân loại đầu vào một cách thích hợp và tạo ra xác suất từ 0 đến 1.
* Lớp phi tuyến tính ReLu: vì tích chập là một phép toán tuyến tính, các lớp tuyến tính thường được đặt trực tiếp sau lớp tích chập để đưa tính phi tuyến tính vào bản đồ kích hoạt. Đơn vị tuyến tính chỉnh lưu (ReLU) tính hàm . Nói cách khác, kích hoạt chỉ đơn giản là ngưỡng bằng không. Đồng thời, lớp này cũng không làm thay đổi cả kích thước của ảnh và thông số. Thông thường, lớp này có đầu ra là ảnh có kích thước hoàn toàn giống với kích thước ảnh đầu vào, chỉ khác là không tồn tại giá trị âm.

1. Cấu trúc của CNN

# THIẾT KẾ VÀ THI CÔNG

## Sơ đồ khối của hệ thống phân loại âm thanh phổi

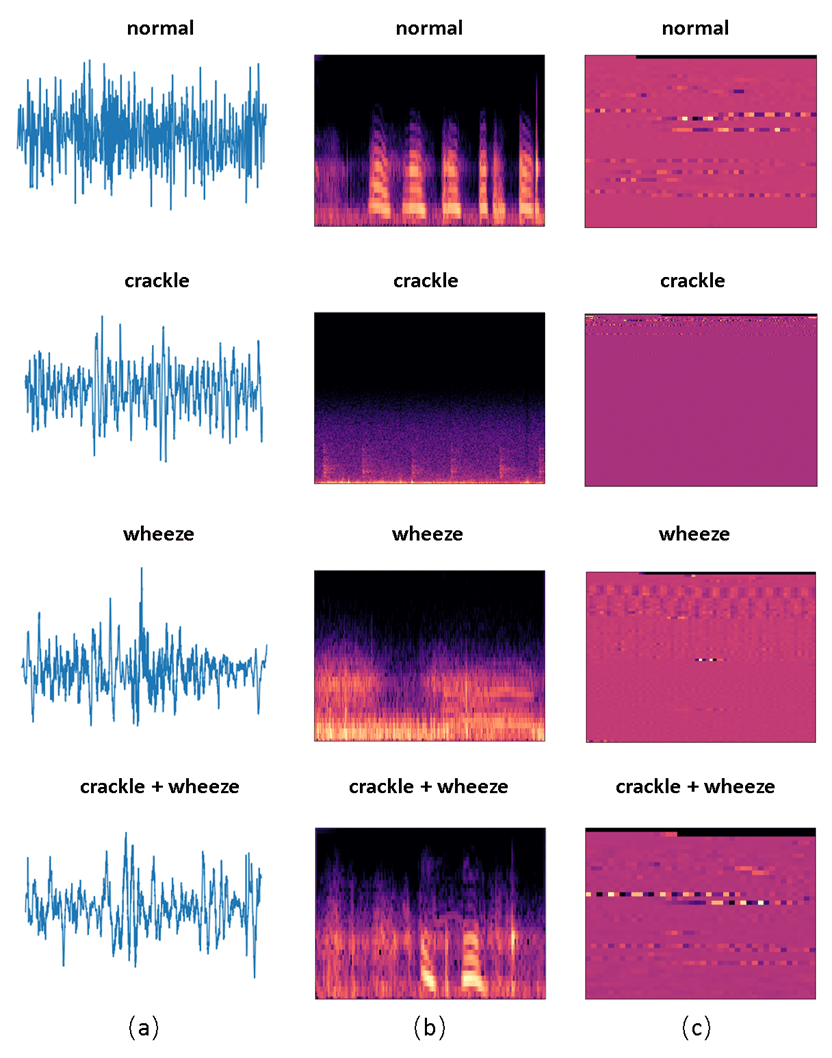
1. Sơ đồ khối của hệ thống

### Khối âm thanh phổi

Âm thanh phổi đã được sử dụng như một công cụ chẩn đoán trong nhiều thế kỷ. Tính hữu ích của việc nghe âm thanh phổi hoặc thính chẩn phổi là một trong những phương pháp phổ biến được sử dụng rộng rãi để đánh giá các bệnh hô hấp. Âm thanh phổi thường được chia làm hai nhóm “âm thanh phổi bình thường” có tần số từ 100 -1000Hz, không có bất kỳ tín hiệu rời rạc nào và “âm thanh phổi bất thường”. Âm thanh phổi bất thường bao gồm âm thanh phổi bất thường liên tục và âm thanh phổi bất thường không liên tục.

Về âm thanh phổi bất thường liên tục gồm có Wheezes là loại âm thanh có vực cao, tần số trên 400Hz và Rhonchi là loại âm thanh có vực thấp, tần số khoảng 200Hz hoặc ít hơn. Bên cạnh đó, Wheezes có thể là đơn âm nếu nó chứa một tần số duy nhất hoặc đa âm khi tần số được cảm nhận đồng thời. Các bệnh liên quan đến âm thanh liên tục là hen suyễn, viêm phổi và viêm phế quản.

Về âm thanh phổi bất thường không liên tục gồm có Crackles có tần số dao động trong khoảng 150 – 2000Hz thường được xác định bởi miền thời gian như chiều rộng lệch của ban đầu (IDW – initial deflection width) và thời gian của hai chu kì (2CD – two-cycle duration). Theo ATS (American Thoracic Society), thời lượng trung bình của IDW và 2CD của Crackles tốt lần lượt là 0.7 và 5ms, và Crackles thô lần lượt là 5 và 10ms. Crackles tốt ở những bệnh viêm phổi, xơ phổi và suy tim sung huyết (CHF).

1. Hình Phân tích phổi âm thanh giữa phổi bình thường, Crackle và Wheeze. (a) Wavelet ban đầu. (b) Phổ STFT của âm thanh phổi. (c) Ma trận phân tích sóng của âm thanh phổi.

### Khối xử lý tín hiệu

Đây là một bước tiền xử lý, tín hiệu thu vào được chuẩn bị cho quá trình xử lý bằng cách thay đổi tốc độ lấy mẫu và chuẩn hóa biên độ với mục đích đồng nhất hóa các âm thanh được sử dụng cho quá trình phân tích. Ở đây tất cả âm thanh được chuyển về 16 bit và có tần số 8000Hz. Quá trình chuẩn hóa dữ liệu bao gồm:

Trong đó, S(n) là tín hiệu đầu vào và N là chiều dài của tín hiệu.

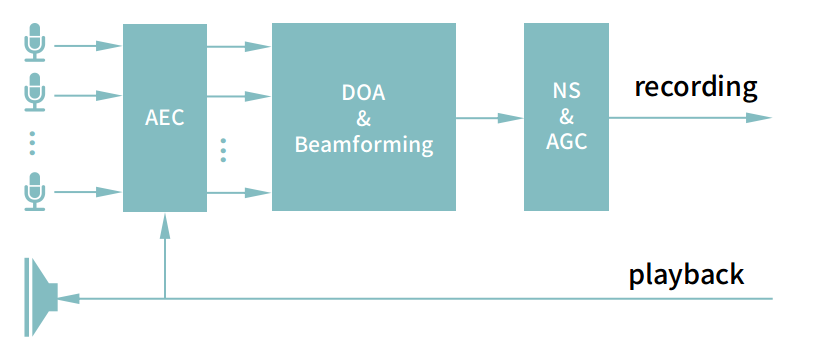
### Khối trích xuất đặc trưng

Việc trích xuất đặc trưng của đối tượng liên quan đến việc giảm số lượng tài nguyên cần thiết để xử lý mà không làm mất thông tin quan trọng hoặc có liên quan cần cho việc phân loại. Ngoài ra, việc giảm dữ liệu tạo điều kiện thuận lợi cho tốc độ học và các bước tổng quát hóa trong quá trình học máy.

### Khối phân loại âm thanh

Phân loại âm thanh dựa trên mô hình CNN với các lớp: lớp tích chập, lớp lấy mẫu và lớp kết nối đầy đủ. Xác suất sự kiện cho tất cả các hình anh trong phân đoạn âm thanh được tích lũy và sau đó sự kiện âm thanh có xác suất tích lũy cao nhất được xác định là kết quả phân loại.

## Sơ đồ khối của hướng âm thanh



1. Sơ đồ khối hướng âm thanh

### Bộ lọc hệ thống AEC (Acoustic Echo Cancellation)

Khử tiếng vọng âm thanh được thiết kế để ngăn chặn tiêng vang và âm thanh không mong muốn được thêm vào từ tín hiệu đi qua không gian âm thanh. Nó dùng bộ lọc FIR và thuật toán sẽ được điều chỉnh liên tục để mô hình hóa đường dẫn âm thanh. Ngoài ra khối AEC cũng tính toán tín hiệu dư ở âm thanh phi tuyến.

Hầu hết AEC được tối ưu hóa cho giọng nói, sử dụng băng thông thấp để giảm thiểu tài nguyên xử lý. Có ba tiêu chuẩn băng thông được dùng rộng rãi: băng tần hẹp tốc độ lấy mẫu là 8 kHz (băng thông 4 kHz), băng tần rộng tốc độ lấy mẫu 16 kHz (băng thông 8 kHz) và băng tần siêu rộng tốc độ lấy mẫu 32 kHz (băng thông 16 kHz).

### DOA (Direction of Arrival)

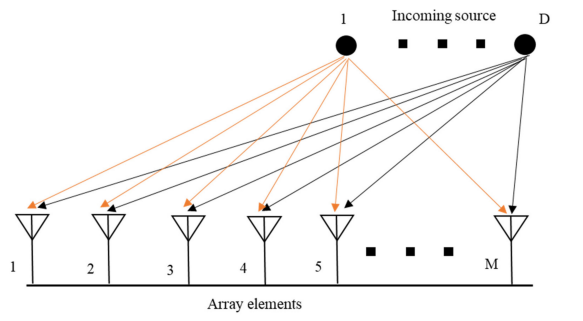
Xác định hướng đến đề cập đến quá trình truy xuất thông tin hướng của một số sóng điện từ hoặc sóng âm, ảnh hưởng đến cảm biến hoặc mảng ăng-ten. Ước tính DOA được sử dụng để định vị và theo dõi nhiều nguồn tín hiệu.

Chuyển sang các thuật toán thực tế để xác định các hướng đến. Mô hình là sự cố tín hiệu M trên mảng, bị nhiễu bởi tiếng ồn, tức là:

Do đó mục tiêu ước tính là , m=1,…M. Cách dễ nhất để ước tính các góc là thông qua mối tương quan. Chúng ta biết rằng bất đẳng thức Cauchy-Schwarz, như một hàm của tối đa tại . Do đó, biểu đồ phương pháp tương quan so với trong đó:

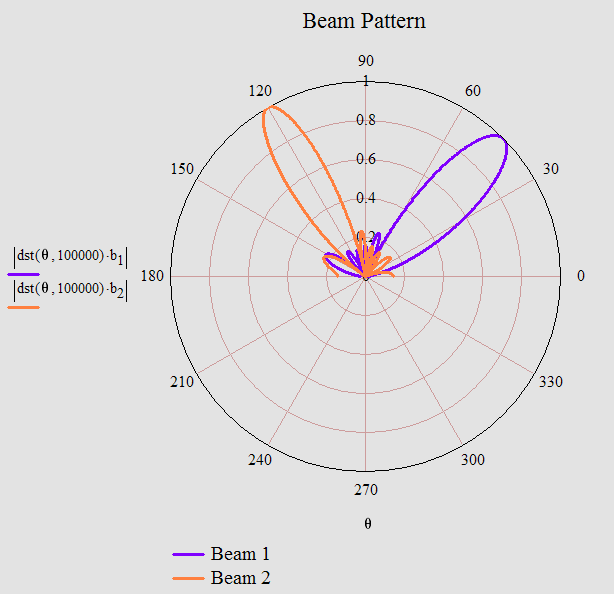
là ước tính không thích ứng của phổ dữ liệu đến. Các đỉnh lớn nhất trong biểu đồ này là các ước tính hướng đến.

Trong trường hợp của mảng tuyến tính, mảng phân bố đều, vecto điều khiển tương đương với các hệ số Fourier, là mối tương quan trong eqn. tương đương với DFT của vecto dữ liệu x.



1. Ước tính DOA

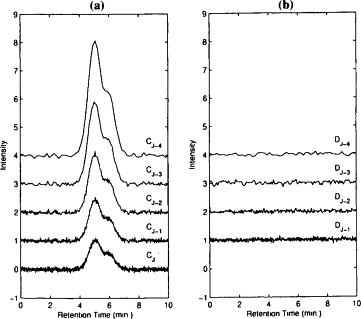
### Beamforming

Beamforming là một kỹ thuật tập trung tín hiệu không dây vào một thiết bị thu cụ thể, thay vì để tín hiệu lan truyền theo mọi hướng. Để tâp trung tín hiệu theo một hướng cụ thể, để tạo thành một chùm năng lượng điện từ nhắm được mục tiêu, nhiều ăng-ten ở gần nhau sẽ phát cùng một tín hiệu vào những thời điểm hơi khác nhau. Các sóng chồng chéo sẽ tạo ra nhiễu, ở một số khu vực mang tính xây dựng (làm cho tín hiệu mạnh hơn) và ở các khu vực mang tính phá hủy (làm cho tín hiệu yếu hơn hoặc không thể phát hiện được). Khi được thực hiện chính xác, quá trình tạo chùm tia này sẽ tập trung tín hiệu theo một hướng cụ thể.

1. Hai mẫu chùm tia

### NS (Noise Suppression)

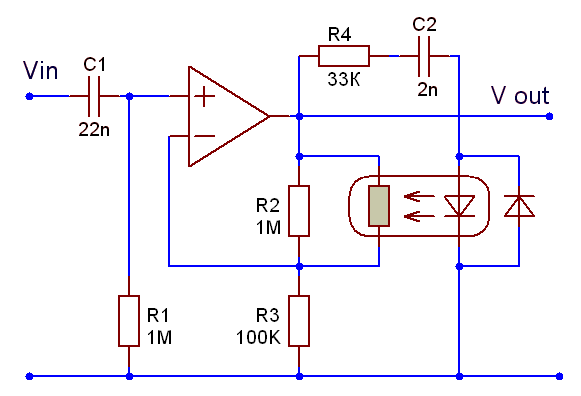
Khử nhiễu là một kỹ thuật rất phổ biến trong xử lý dữ liệu sắc ký. Nó nhằm mục đích tăng cường tín hiệu phân tích để đưa ra tỷ lệ nhiễu tín hiệu cao hơn. Ngày nay, nhiều dụng cụ sắc ký được kiểm soát bởi các máy tính và nó đã trở thành một thông lệ để giảm tiếng ồn bằng cách sử dụng các phương pháp xử lý kỹ thuật số như lọc. Theo truyền thống, các nhà khoa học áp dụng các bộ lọc Savitzky-Golay, Fourier và Kalman để xử lý tín hiệu.



1. Tín hiệu vào (a) và tín hiệu sau khi lọc nhiễu (b)

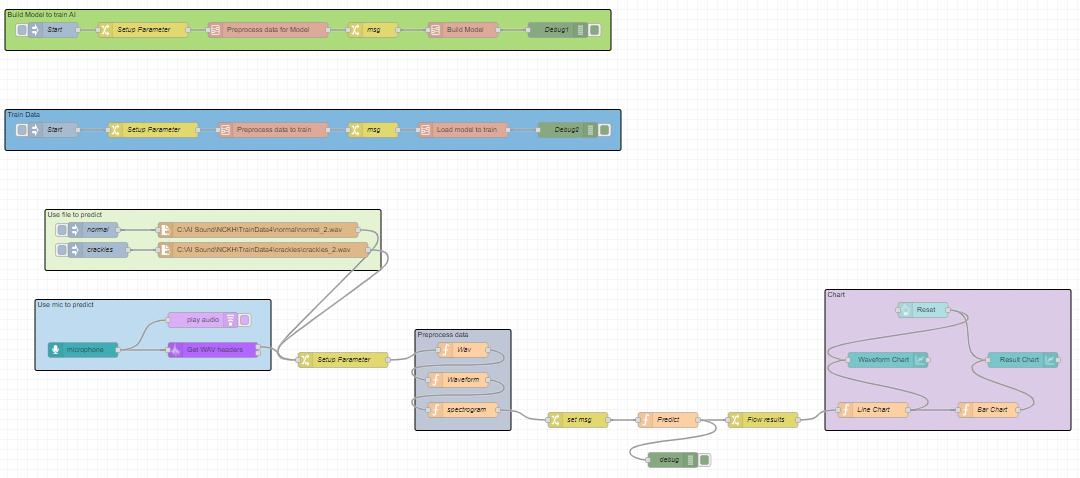
### AGC (Automatic Gain Control)

Mạch điều khiển độ lợi tự động là một mạch được thiết kế để duy trì mức tín hiệu đầu ra không đổi sau khi khuếch đại, không quan tâm đến sự thay đổi tín hiệu ở đầu vào của bộ khuếch đại hoặc hệ thống. Trong điều kiện môi trường ít bị suy heo, mạch AGC sẽ làm giảm độ lợi do bộ khuếch đại cung cấp để nó không làm hỏng mạch thu.

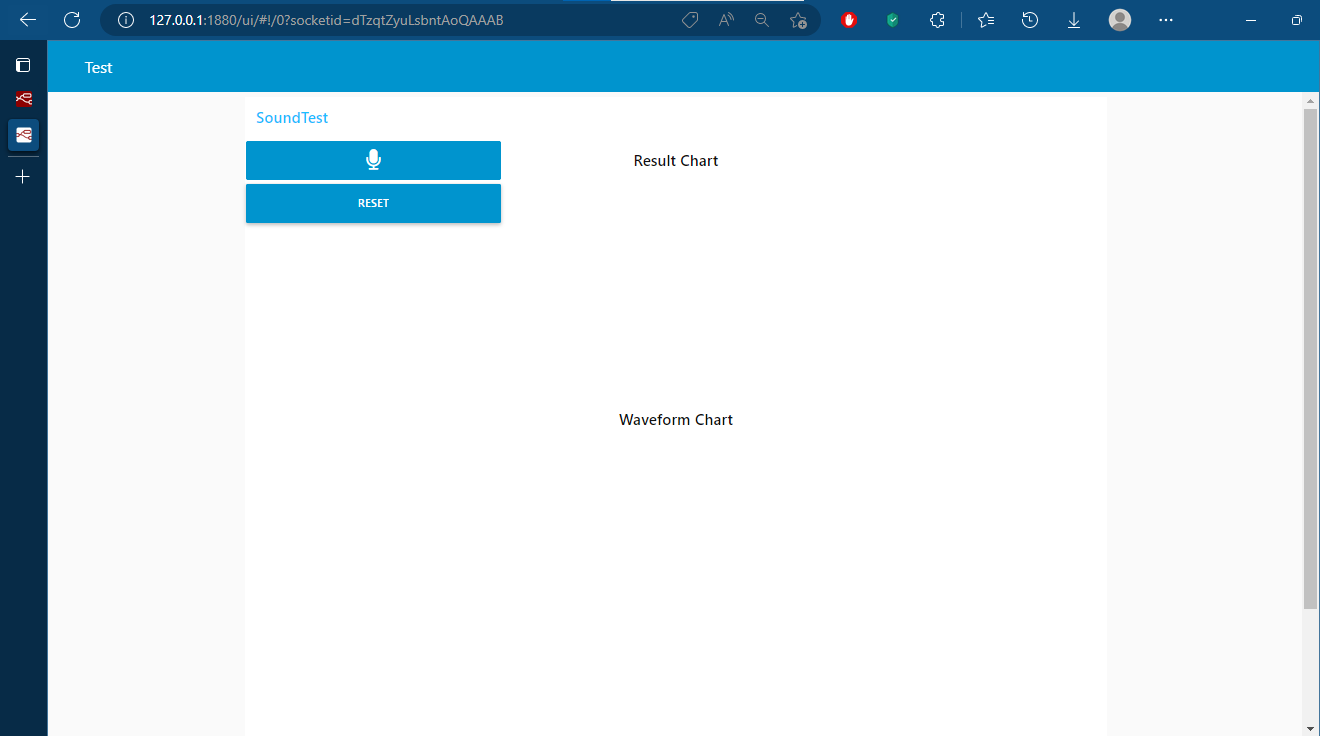


1. Sơ đồ của một AGC được sử dụng trong mạng điện thoại analog

## Thiết kế giao diện



1. Giải thuật trên node red



1. Giao diện dashboard của node red

# GIẢI THUẬT VÀ ĐIỀU KHIỂN

## Hoạt động của hệ thống

### Phân loại âm thanh phổi

Khi bắt đầu ta đưa một tệp âm thanh vào hệ thống. Sau đó, thống qua bước tiền xử lý để xử lý các dữ liệu thô bằng cách thay đổi tốc độ lấy mẫu. Tiếp theo, âm thanh được xử lý thông qua quá trình chuẩn hóa biên độ thu được dữ liệu được chuẩn hóa. Tiếp đến, dữ liệu thông qua biến đổi Fourier nhanh làm tăng khả năng, tốc độ tính toán của DFT. Qua đó đến bước phân loại, cho ra được kết quả phân loại cuối cùng và kết thúc.

### Hướng âm thanh

Âm thanh được ghi lại từ micro sau đó được gửi đến đường dẫn âm thanh (một bộ khuếch đại). AEC dưa trên bộ lọc FIR, đầu ra của bộ lọc sau đó được trừ vào tín hiệu đường dẫn âm thanh để tạo ra đầu ra tín hiệu “sạch” với phần tuyến tính của tiếng vang âm thanh phần lớn bị loại bỏ. Tiếp theo, tín hiệu sẽ được truyền đến DOA và beamforming để ước tính hướng nhận được, tập trung tín hiệu từ một hướng đến mạch. Tiếp đến, tín hiệu nhận được từ một hướng xác định được đưa đến để khử nhiễu do các tín hiệu gây ra và tính toán độ lợi tự động để khuếch đại tín hiệu và đưa vào hệ thống phân loại âm thanh để xử lý.

## Lưu đồ giải thuật

### Phân loại âm thanh phổi

1. Lưu đồ giải thuật của hệ thống phân loại âm thanh phổi

### Hướng âm thanh



1. Lưu đồ giải thuật hướng âm thanh

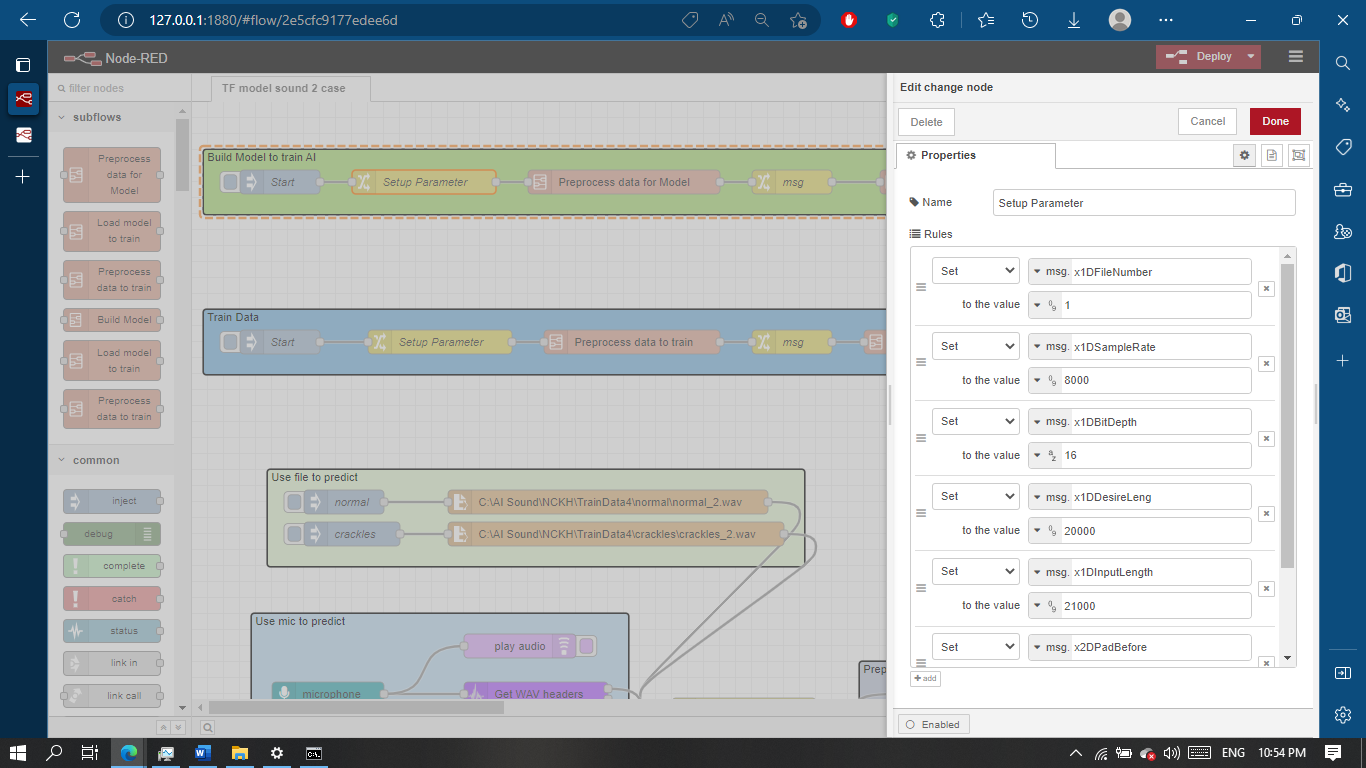
# THỰC NGHIỆM

## Tiến trình thực nghiệm

### Phân loại âm thanh phổi

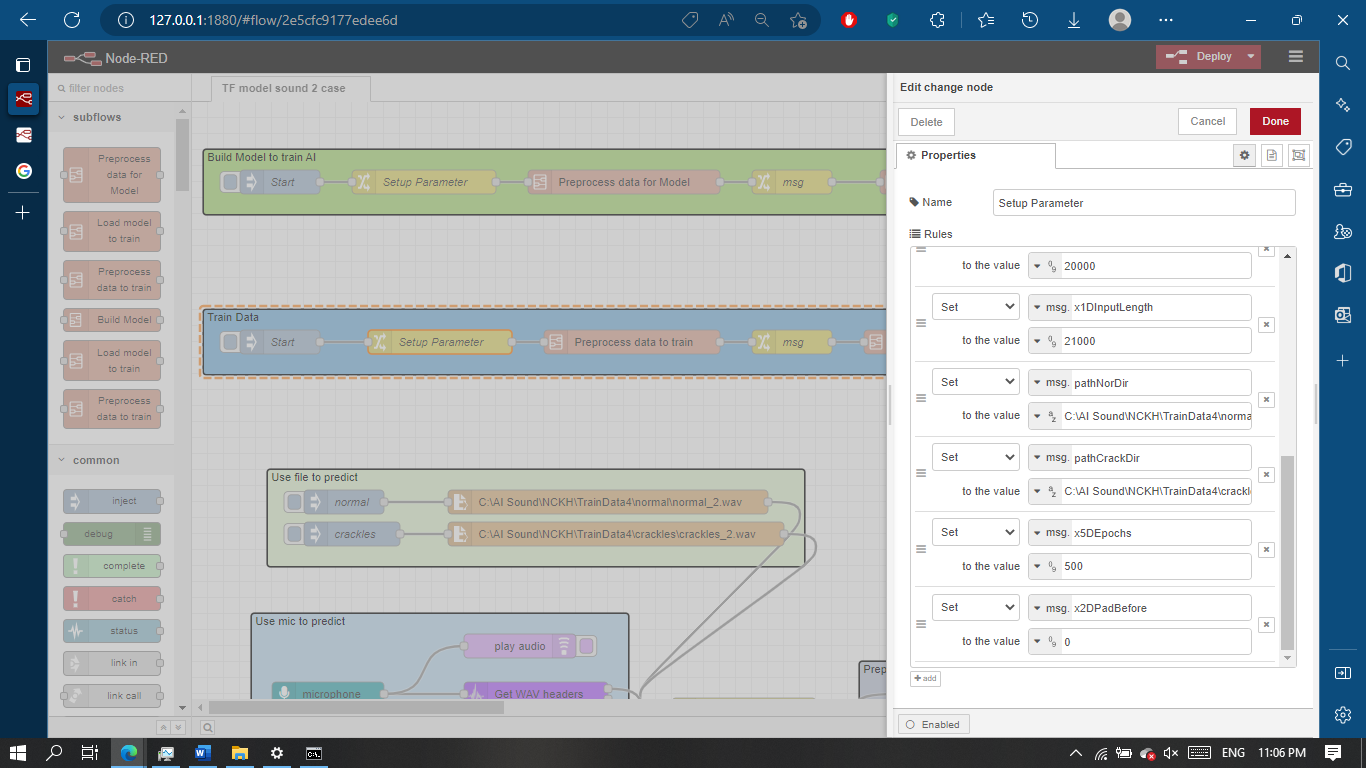
Bước 1: Truy cập node red bằng nhập lệnh node-red trong cmd (http://127.0.0.1:1880/)

Bước 2: Truy cập dashboard của node red (<http://127.0.0.1:1880/ui>)

Bước 3: Chỉnh sửa các thông số, đường dẫn đến các tệp tại Setup Parameter tại Build Model to Train AI. Cài đặt 16 bit (x1DbitDepth), tần số 8000Hz (x1DsampleRate), độ dài mong muốn 20000 (x1DdesireLeng), độ dài ngõ vào 21000 (x1DinputLength), dẫn đường dẫn vào thư mục chứa file normal để xây dựng model.

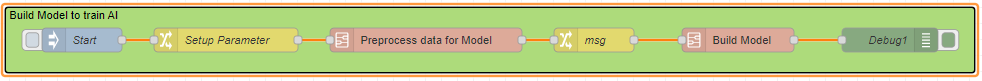
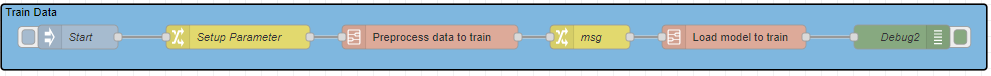
1. Setup Parameter trong block Build Model to Train AI

Bước 4: Chỉnh sửa các thông số, đường dẫn đến các tệp tại Setup Parameter tại Train data. Cài đặt 16 bit (x1DbitDepth), tần số 8000Hz (x1DsampleRate), độ dài mong muốn 20000 (x1DdesireLeng), độ dài ngõ vào 21000 (x1DinputLength), dẫn đường dẫn vào thư mục chứa file normal (pathNorDir), dẫn đường dẫn vào thư mục crack (pathCrackDir) và thiết lập số lần train 500 (x5DEpochs) để bắt đầu train model.

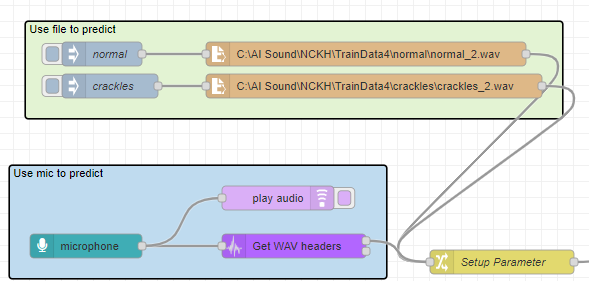
1. Setup Parameter trong block Train Data

Bước 5: Sau khi chỉnh sửa thì lưu lại bằng cách nhấn Deploy

Bước 6: Nhấn vào Start tại block Build Model to Train AI để xây dụng mô hình và Start tại block Train data để thực hiện đào tạo



1. Block Build Model to Train AI và block Train Data

Bước 7: Sau khi hoàn tất train, bắt đầu dự đoán kết quả. Có 2 cách để đưa âm thanh vào là sử dụng file có sẵn hoặc là sử dụng micro thu trực tiếp qua mạch để dự đoán và phân loại âm thanh.

1. Các cách đưa dữ liệu vào

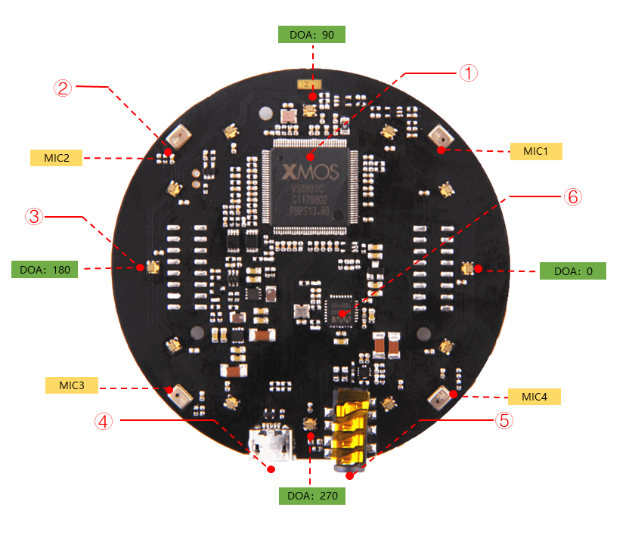
### Hướng âm thanh



1. Mạch thu âm thanh thực tế

Mạch ReSpeaker Mic Array v2.0 bao gồm :

1. XMOS XVF-3000: Nó tích hợp các thuật toán DSP tiên tiến bao gồm khử tiếng vọng âm thanh (AEC), beamforming, âm dội, khử tiếng ồn (NS) và kiểm soát độ lợi (AGC).
2. Micrô kỹ thuật số: MP34DT01-M là micrô MEMS kỹ thuật số siêu nhỏ gọn, công suất thấp, đa hướng, được chế tạo với phần tử cảm biến điện dung và giao diện IC.
3. Đèn Led RGB: ba màu.
4. Cổng USB: Cung cấp nguồn điện và điều khiển mảng micrô.
5. Giắc cắm tai nghe 3.5 mm: Đầu ra âm thanh, Chúng ta có thể cắm loa hoạt động hoặc tai nghe vào cổng này.
6. WM8960 là codec âm thanh nổi công suất thấp có trình điều khiển loa Class D để cung cấp 1 W mỗi kênh thành tải 8 W.

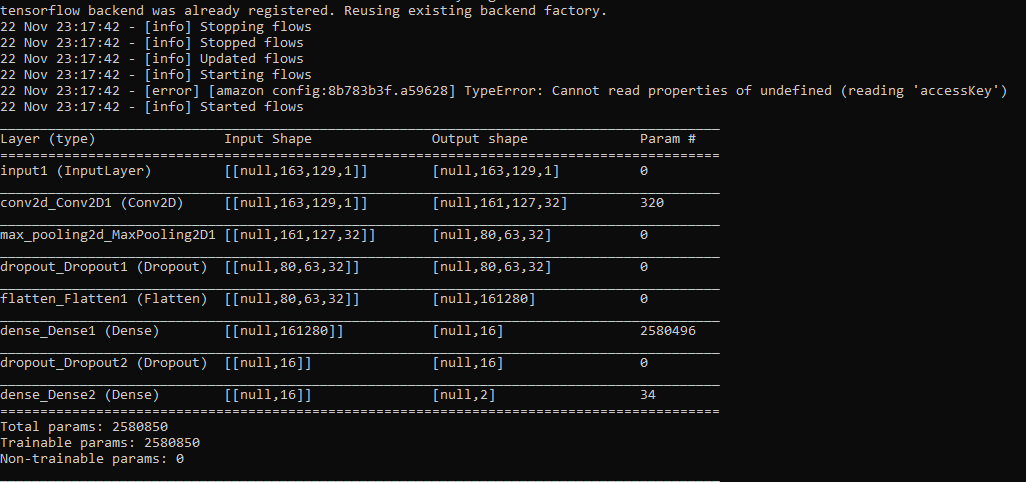


1. Chi tiết mạch thu âm thanh

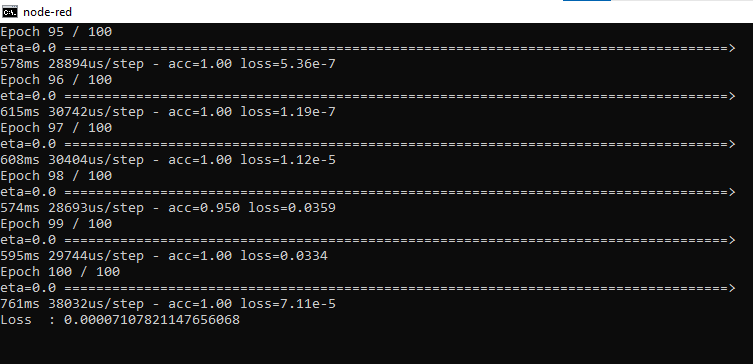
## Kết quả thực nghiệm

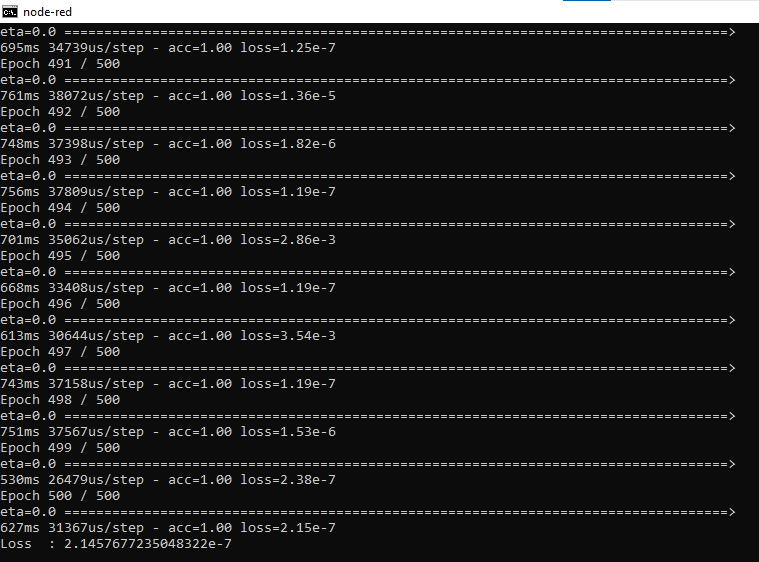
### Kết quả đào tạo

Sau khi build model ta được:

1. Model thu được sau khi xây dựng

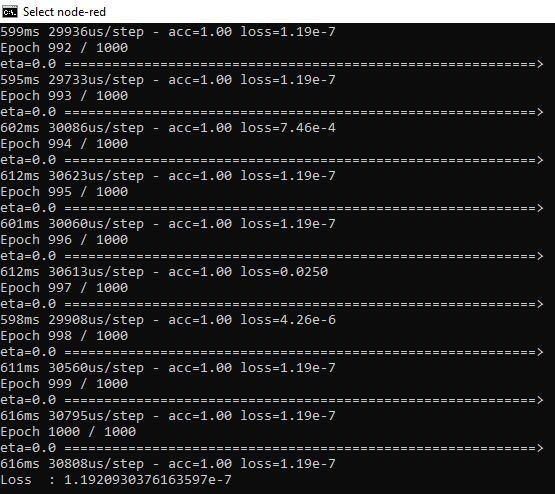
Sau khi học với 100 epoch thì kết quả cho ra được độ chính xác (acc) là 1.00 và độ sai lệch (loss)

1. Thông số cuối cùng của 100 epoch

Sau khi học với 500 epoch thì kết quả cho ra được độ chính xác (acc) là 1.00 và độ sai lệch (loss)

1. Thông số cuối cùng của 500 epoch

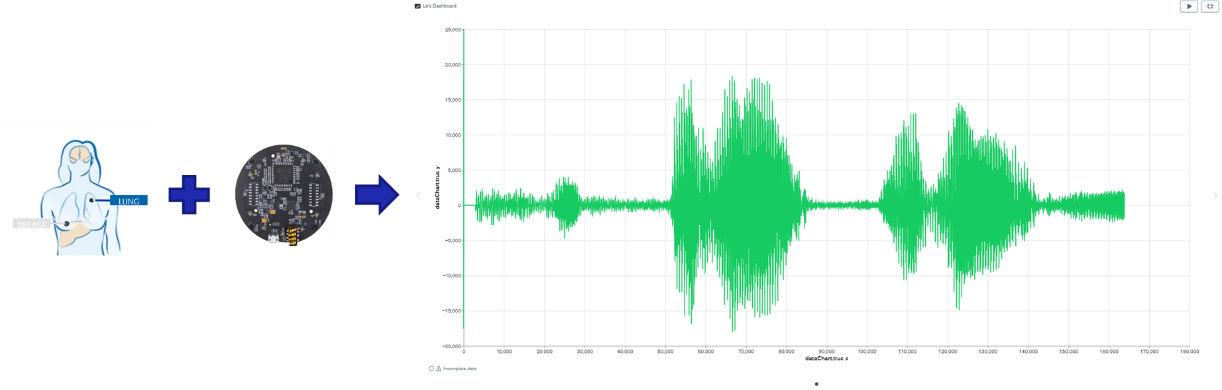
Sau khi học với 1000 epoch thì kết quả cho ra được độ chính xác (acc) là 1.00 và độ sai lệch (loss)



1. Thông số cuối cùng của 1000 epoch

### Kết quả hướng âm thanh

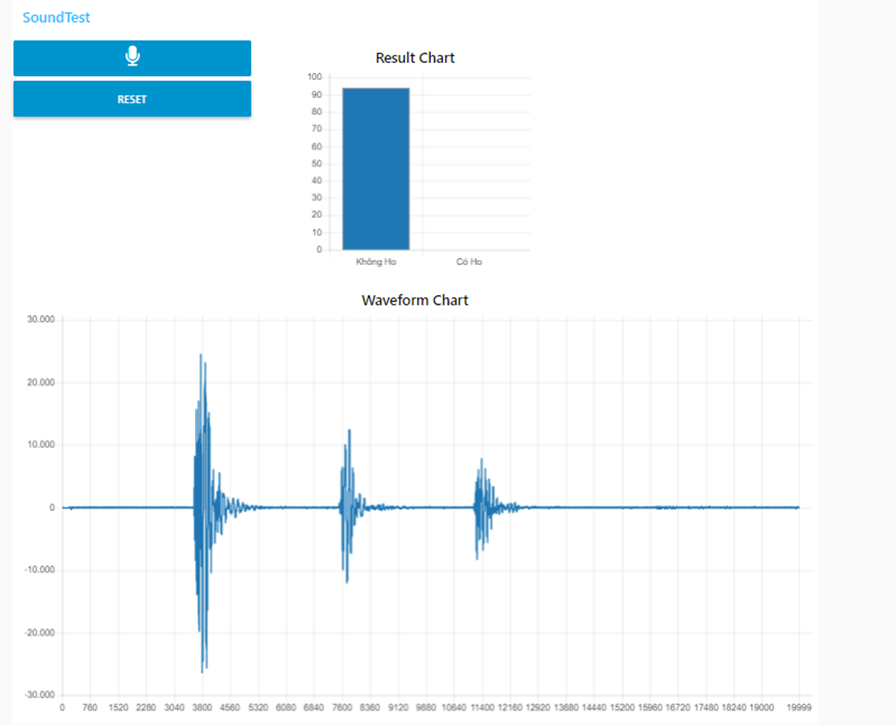
Sử dụng mạch thu âm thanh



1. Âm thanh thu được sau khi được xử lý thông qua mạch

### Kết quả dự đoán

Dùng model 100 epoch ta thu được:



1. Tín hiệu đầu vào micro (người khỏe mạnh) cho ra kết quả không ho



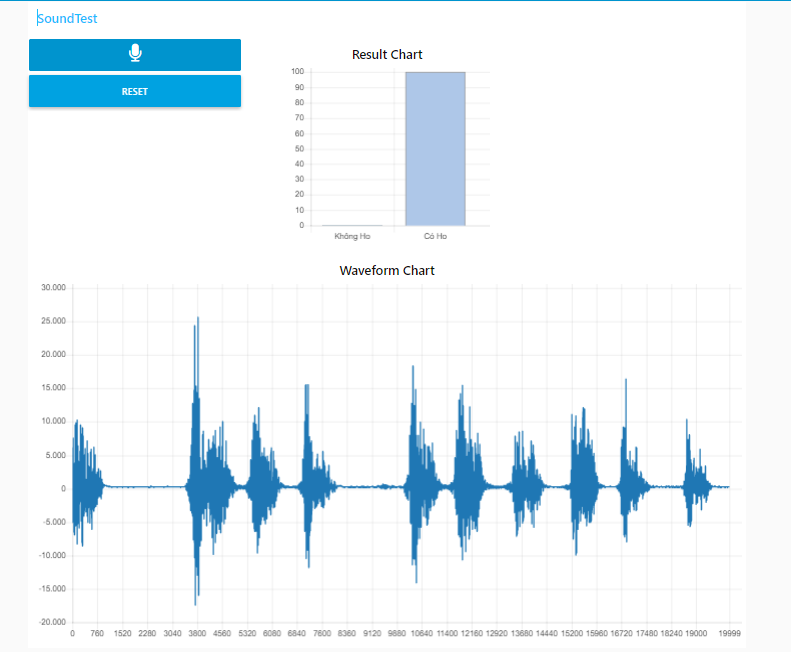
1. Tín hiệu đầu ra của micro (không ho)



1. Tín hiệu đầu vào file normal (2) cho ra kết quả không ho



1. Tín hiệu đầu ra của file normal

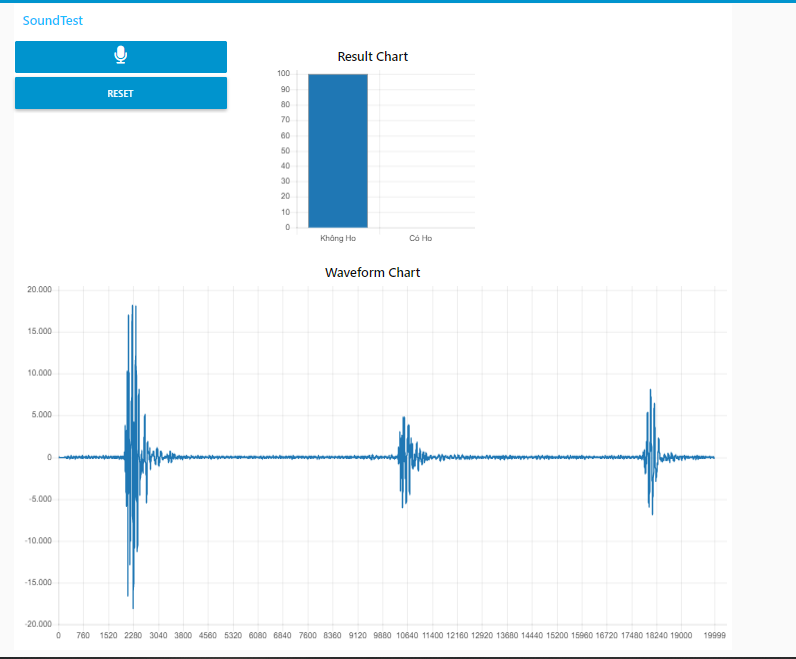


1. Tín hiệu đầu vào file crack (2) cho ra kết quả có ho



1. Tín hiệu đầu ra của file crack

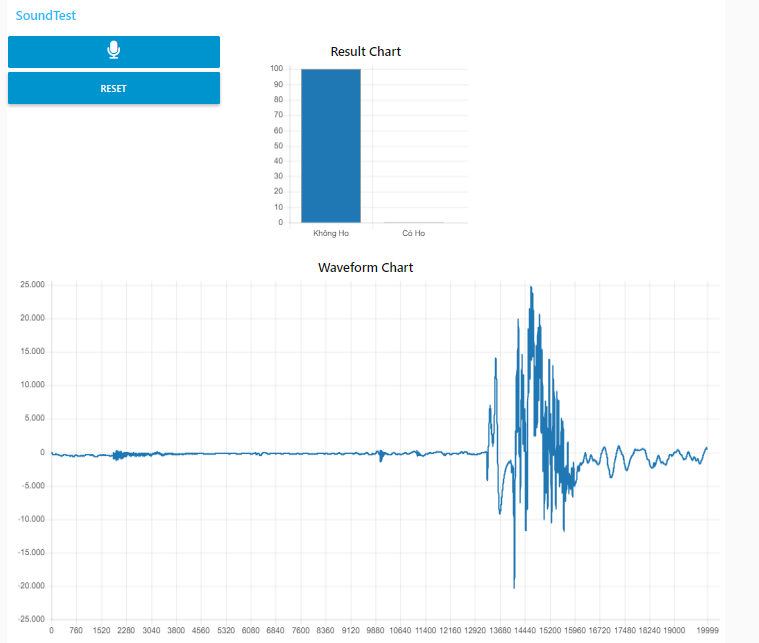
Dùng model 500 epoch ta thu được:



1. Tín hiệu đầu vào micro (người khỏe mạnh) cho ra kết quả không ho



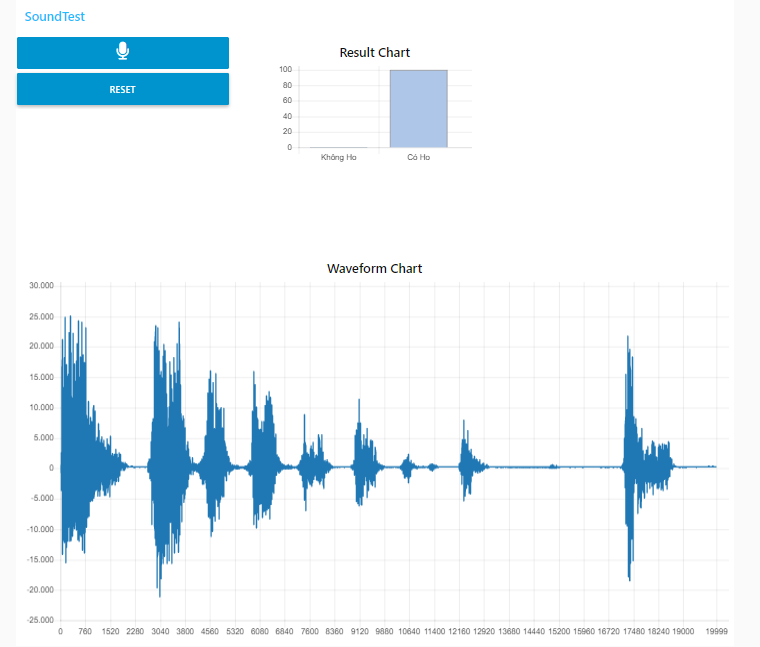
1. Tín hiệu đầu ra của micro (không ho)



1. Tín hiệu đầu vào file normal (5) cho ra kết quả không ho



1. Tín hiệu đầu ra của file normal

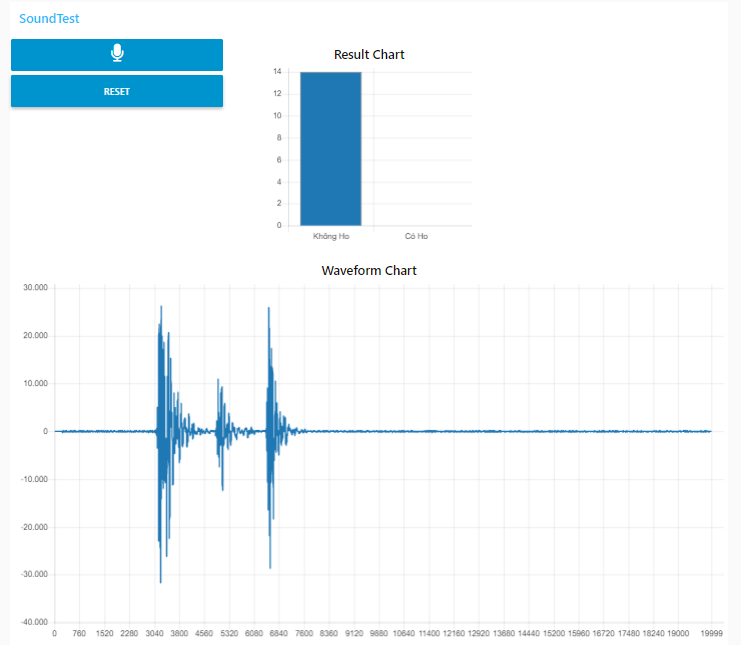


1. Tín hiệu đầu vào file crack (5) cho ra kết quả có ho



1. Tín hiệu đầu ra của file crack

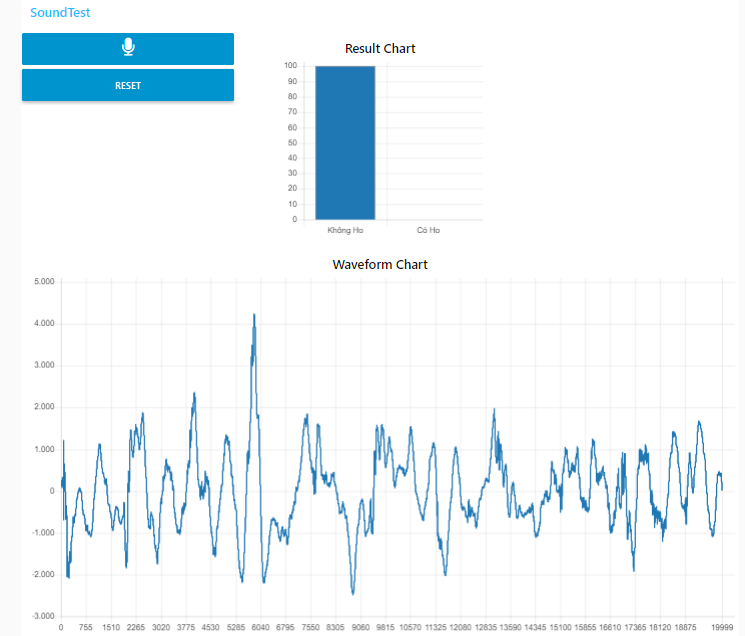
Dùng model 1000 epoch ta thu được:



1. Tín hiệu đầu vào micro (người khỏe mạnh) cho ra kết quả không ho



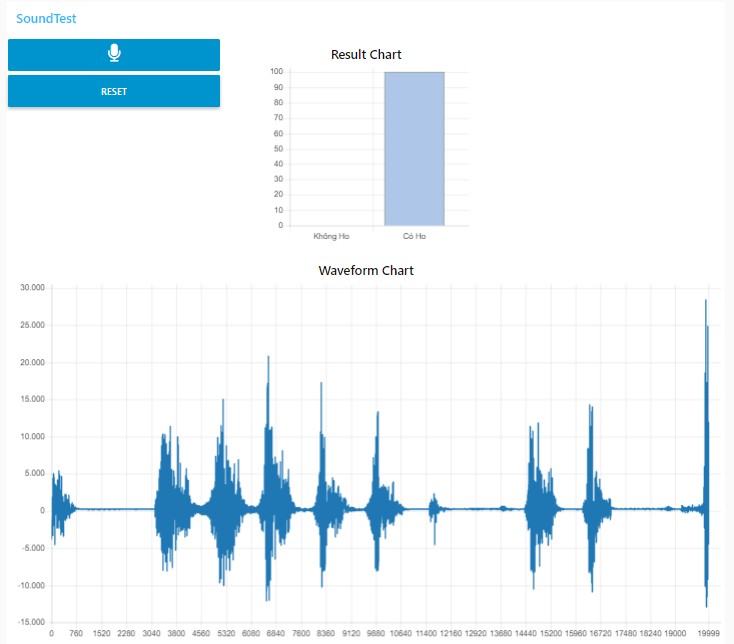
1. Tín hiệu đầu ra của micro (không ho)



1. Tín hiệu đầu vào file normal (3) cho ra kết quả không ho



1. Tín hiệu đầu ra của file normal



1. Tín hiệu đầu vào file crack (3) cho ra kết quả có ho



1. Tín hiệu đầu ra của file crack

## Kết luận thực nghiệm

* Thông qua các lần train có số epoch khác nhau thì số epoch càng lớn thit tỷ lệ sai số càng ít và độ chính xác càng cao.
* Cho ra được giao diện của laptop.
* Nhận dạng được hướng âm thanh.
* Phân loại được giữa âm thanh phổi của người khỏe mạnh và âm thanh phổi bất thường.

# KẾT LUẬN

## Ưu điểm

* Tốc độ học nhanh và chính xác khi sử dụng tensorflow.
* Phân loại được giữa phổi bình thường và phổi bất thường.
* Tính ứng dụng cao trong y học.

## Nhược điểm

* Âm thanh đưa vào còn nhiễu nhiều.
* Khó phân biệt được tiếng ho với các tiếng tương tự khác (như tiếng gõ bàn,…)

## Hướng phát triển

* Dùng phần mềm để loại bỏ nhiễu tại tín hiệu đưa vào mà không cần đến phần cứng.

# TÀI LIỆU THAM KHẢO

1. Anthony Reinette, M. C. (2017, August 24). *Benchmarking Microphone Arrays: ReSpeaker, Conexant, MicroSemi AcuEdge, Matrix Creator, MiniDSP, PlayStation Eye*. Retrieved from Medium: https://medium.com/snips-ai/benchmarking-microphone-arrays-respeaker-conexant-microsemi-acuedge-matrix-creator-minidsp-950de8876fda?fbclid=IwAR3xKWaHTbI1MFKSNd8HX6NjpDtSOPuxh23XowBlteTS3Rkjheiu0gkrcIA
2. Berlin, J. (2021, November 30). *Acoustic Echo Cancellation*. Retrieved from Wiki analog devices: https://wiki.analog.com/resources/tools-software/sigmastudio/toolbox/adialgorithms/aec
3. Bryan, N. (2016). *B3. Short Time Fourier Transform.* California: Naval Postgraduate School.
4. Dhabale, A. (2018). *Direction Of Arrival (DOA) Estimation Using Array Signal Processing.* CALIFORNIA: Peer reviewed.
5. FATIH DEMIR, A. M. (2017). *Classification of lung sounds with CNN model using parallel pooling structure.* IEEE Access.
6. Gasmi, A. (2022). *What is Fast Fourier Transform?*
7. Grønnesby, M. (2016). *Automated Lung Sound Analysis.* Norway: Master Thesis in Computer Science.
8. Hangting Chen, P. Z. (2018). *Deep Convolutional Neural Network with Scalogram for Audio Scene Modeling.* China: Interspeech 2018.
9. M.A. Tocchetto, A. B. (2014). *An Embedded Classifier of Lung Sounds based on the Wavelet Packet Transform and ANN.* South Africa: IFAC.
10. Minkyu Lim, D. L.-S.-J.-H. (2018). *Convolutional Neural Network based Audio Event Classification.* South Korea: KSII.
11. Mishra, M. (2020, August 26). *Convolutional Neural Networks, Explained*. Retrieved from Towards Data Science: https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939
12. Shreyasi Datta, A. D. (2017). *Automated Lung Sound Analysis for Detecting Pulmonary Abnormalities.* IEEE.
13. Wesfreid, E. (2013). *STFT time-frequency visualization Application to sound signals.* France: Image Processing On Line.
14. *What is a Sound Spectrum?* (n.d.). Retrieved from University New South Wales: http://www.phys.unsw.edu.au/jw/sound.spectrum.html
15. Yoonjoo Kim, Y. H. (2021). *Respiratory sound classifcation for crackles, wheezes, and rhonchi in the clinical feld using deep learning.* Scientifc Reports .
16. CHƯƠNG TRÌNH HƯỚNG ÂM THANH

import json

from tuning import Tuning

import usb.core

import usb.util

import time

dev = usb.core.find(idVendor=0x2886, idProduct=0x0018)

if dev:

Mic\_tuning = Tuning(dev)

print Mic\_tuning.direction

while True:

try:

print Mic\_tuning.direction

time.sleep(1)

dataDOA = {'angle' : Mic\_tuning.direction}

json\_string = json.dumps(dataDOA)

with open('json\_DOA.json', 'w') as outfile:

outfile.write(json\_string)

except KeyboardInterrupt:

break