

DOI:10.22144/ctu.jvn.2022.003

KHẢO SÁT KỸ THUẬT HỌC SÂU TRÊN BÀI TOÁN CHẨN ĐOÁN HƯ HỒNG ĐỘNG CƠ ĐIỆN DỰA TRÊN TIẾNG ÒN VẬN HÀNH

Nguyễn Chí Ngôn^{1*}, Lê Thanh Tú², Lương Hoàng Vĩnh Thuận³ và Nguyễn Chánh Nghiệm¹

¹Trường Đại học Cần Thơ

²Trường Cao đẳng Long An

³Trường Cao đẳng nghề An Giang

*Người chịu trách nhiệm về bài viết: Nguyễn Chí Ngôn (email: ncngon@ctu.edu.vn)

Thông tin chung:

Ngày nhận bài: 19/08/2021

Ngày nhận bài sửa: 11/09/2021

Ngày duyệt đăng: 26/02/2022

Title:

Investigation of deep learning algorithm applied on induction motor fault diagnosis based on operation sound

Từ khóa:

Biến đổi wavelet, chẩn đoán lỗi động cơ, mạng GoogLeNet, mạng neuron tích chập, phổ ảnh tín hiệu

Keywords:

Convolutional neural networks, GoogLeNet, motor fault diagnosis, scalogram, wavelet transform

ABSTRACT

Early detection of induction motor failure plays an important role in limiting disruption to industrial production. Sensor-based measurement methods are highly reliable, but the installation of the equipment is time consuming and costly. Building a smartphone's application to diagnose electric motor problems is a research direction that attracts many groups. This paper proposes to investigate the ability to train and diagnose electric motor faults based on the principle of recognizing scalogram images of motor's operation sounds, using a deep learning neural network. The audio signals are noise-filtered, amplitude normalized, and scalogram rendering by wavelet transforms. The set of scalogram images is divided into two parts for training and validating the GoogLeNet convolutional neural network. The GoogLeNet is also investigated through changing some basic parameters, in order to determine the best training efficiency. After training, the network is tested on an independent sound signal dataset. The results show that the network is able to identify 3 common motor problems including phase loss, insulating film brush and bearing failure with 94.21% accuracy. The experiment also shows that the development of smartphone's application for early diagnosing electric motor problems is feasible.

TÓM TẮT

Phát hiện sớm sự cố động cơ điện góp phần hạn chế gián đoạn hoạt động sản xuất công nghiệp. Phương pháp đo dùng cảm biến có độ tin cậy cao, song việc lắp đặt mất thời gian và chi phí. Việc xây dựng ứng dụng điện thoại để chẩn đoán sự cố động cơ điện thu hút nhiều nghiên cứu. Bài báo tiến hành khảo sát khả năng chẩn đoán lỗi động cơ điện thông qua nhận diện ảnh phổ tín hiệu âm thanh vận hành dùng mạng neuron học sâu GoogLeNet. Dữ liệu âm thanh được lọc nhiễu, chuẩn hóa biên độ và dựng ảnh phổ bằng phép biến đổi wavelet. Tập ảnh phổ được dùng để huấn luyện và kiểm tra mạng. Mạng GoogLeNet cũng được khảo sát hiệu quả huấn luyện thông qua việc thay đổi các tham số cơ bản. Sau đó, mạng được kiểm tra trên tập dữ liệu độc lập. Kết quả cho thấy mạng nhận diện 3 sự cố thông dụng, gồm mất pha, cọ phim và hỏng bạc đạn, với tỷ lệ chính xác đạt 94,21%. Thí nghiệm cũng cho thấy khả năng phát triển ứng dụng điện thoại là khả thi.

1. GIỚI THIỆU

Động cơ điện nói chung và động cơ cảm ứng (induction motor) nói riêng có vai trò quan trọng trong công nghiệp. Tuy nhiên, với cường độ hoạt động cao, động cơ điện thường xảy ra một số hư hỏng, như lỗi bạc đạn (vòng bi), lỗi stator và lỗi rotor (Bonnett et al., 1992; Henao et al., 2014). Việc phát hiện sớm sự cố cho động cơ điện là một yêu cầu hết sức cần thiết và quan trọng, góp phần hạn chế việc gián đoạn trong sản xuất công nghiệp.

Nghiên cứu chẩn đoán lỗi động cơ điện được nhiều nhà khoa học quan tâm. Nổi bật là các phương pháp phân tích rung động (Kim et al., 2014; Li et al., 2017), phân tích dòng điện (Ciszewski et al., 2016; Yang et al., 2016; Glowacz et al., 2017) và phân tích nhiệt độ (Guo et al., 2012). Các phương pháp này áp dụng kỹ thuật đo điện thông qua các cảm biến được bố trí xung quanh động cơ điện. Kỹ thuật này đôi lúc gặp khó khăn trong việc triển khai cụ thể đo trong môi trường công nghiệp. Từ thực tế đó, các kỹ thuật đo không điện đã được phát triển. Chẳng hạn, phương pháp phân tích ảnh nhiệt đã được áp dụng (Glowacz et al., 2017). Tuy nhiên, kỹ thuật này có nhược điểm cơ bản là chi phí cao, cần làm nóng động cơ, xử lý dữ liệu ảnh nhiệt chậm và bố trí máy ảnh nhiệt khó khăn. Phương pháp phân tích âm thanh của động cơ (Singh et al., 2017; Glowacz et al., 2018) cũng được triển khai. Ưu điểm của phương pháp này là áp dụng kỹ thuật không xâm lấn, chi phí thấp và đo tức thời âm thanh phát ra từ động cơ đang vận hành. Tuy nhiên, nhược điểm của phương pháp phân tích âm thanh động cơ là vấn đề xử lý tạp âm trong môi trường công nghiệp và cũng như đòi hỏi mức độ tin cậy cao của giải thuật chẩn đoán.

Việc ứng dụng phép biến đổi wavelet (Ince et al., 2016; Sangeetha et al., 2017) để lọc nhiễu và tạo ảnh phổ 2 chiều đã được phát triển mạnh mẽ. Ngoài ra, nhờ tốc độ của máy tính ngày càng được cải thiện, các mạng neuron tích chập (CNN – convolutional neural networks) cùng với phương pháp học sâu (deep learning) ngày càng được ứng dụng trong nhiều lĩnh vực. So với phương pháp học máy truyền thống, phương pháp học sâu đã thể hiện được nhiều ưu

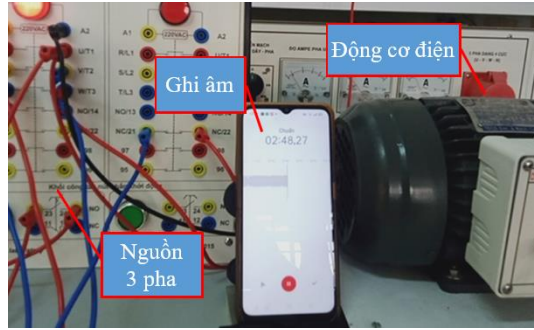
điểm hơn, nhất là trong lĩnh vực dự báo lỗi động cơ (Qi et al., 2017; Wen et al., 2017; Xia et al., 2018). Tuy nhiên, việc áp dụng phép biến đổi wavelet kết hợp với mạng CNN để chẩn đoán lỗi động cơ (Kumar et al., 2020) vẫn còn giới hạn về các công bố khoa học. Ứng với bài toán chẩn đoán lỗi động cơ điện dùng mạng học sâu CNN, thì các tham số của quá trình huấn luyện mạng cần được khởi tạo như thế nào cũng là một câu hỏi chưa được giải đáp thỏa đáng.

Trong nghiên cứu này, phương pháp chẩn đoán lỗi động cơ cảm ứng dựa trên phân tích tín hiệu âm thanh thu được trong lúc vận hành kết hợp với mạng học sâu CNN và kỹ thuật lọc nhiễu, tạo ảnh phổ 2 chiều bằng phép biến đổi wavelet. Mục tiêu chính của nghiên cứu là tiến hành khảo sát ảnh hưởng của một số tham số cơ bản của kỹ thuật học sâu đến quá trình huấn luyện và khả năng nhận diện lỗi động cơ của mạng CNN. Để đạt được mục tiêu này, nghiên cứu đã bố trí thí nghiệm để ghi âm tiếng ồn phát ra trong lúc động cơ vận hành bằng điện thoại di động, tương ứng với các tình huống sự cố, gồm: vỡ bạc đạn, mất pha và cọ phim cách điện. Dữ liệu âm thanh sau đó được tiền xử lý thông qua việc lọc nhiễu bằng phép biến đổi wavelet và chuẩn hóa biên độ tín hiệu. Tập dữ liệu sau tiền xử lý sẽ được biến đổi thành tập ảnh phổ tần số hai chiều, cũng dùng phép biến đổi wavelet. Mô hình mạng CNN của Google (gọi là GoogLeNet, Mishra et al., 2020) được sử dụng và được khảo sát hiệu quả với sự thay đổi của các tham số cơ bản của mạng.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Bố trí thí nghiệm

Thí nghiệm ghi âm tiếng ồn của động cơ điện được thực hiện nhằm thu thập dữ liệu âm thanh phát ra, tương ứng với các trường hợp động cơ mất pha, cọ phim cách điện và vỡ bạc đạn. Thông số của động cơ điện trong thí nghiệm như sau: Động cơ không đồng bộ 3 pha, điện áp định mức 220/380V, dòng định mức 1,2 A, tốc độ 960 vòng/phút, công suất định mức 0,37 kW (0,5 HP) và tần số 50 Hz. Thiết bị ghi âm là điện thoại di động thông minh phổ thông, tần số lấy mẫu 44.100 Hz. Mô hình thí nghiệm được bố trí như Hình 1. Thông tin về dữ liệu thu thập được tóm tắt ở Bảng 1.



Hình 1. Bố trí thí nghiệm

Bảng 1. Dữ liệu thí nghiệm

Thông tin	Tình huống		
	Mất pha	Cọ phim	Hồng bạc đạn
Thời gian ghi âm mỗi tập tin (giây)	120	120	120
Số tập tin âm thanh để huấn luyện mạng	35	35	35
Số tập tin âm thanh độc lập để kiểm thử	4	5	5

Dữ liệu âm thanh ghi được thường bị tác động bởi nhiễu, đồng thời, tùy thuộc vào các tình huống hư hỏng khác nhau mà biên độ tín hiệu âm thanh vận hành thu được cũng có giá trị khác nhau. Vì vậy, khâu tiền xử lý dữ liệu bao gồm lọc nhiễu và chuẩn hóa biên độ tín hiệu về giá trị [-1, 1] sẽ được áp dụng trong phần tiếp theo.

2.2. Lọc nhiễu tín hiệu bằng phép biến đổi wavelet

Phép biến đổi wavelet được đề cập bởi Haar (1910), nhưng nó chỉ phát triển mạnh mẽ trong xử lý tín hiệu từ khi Mallat (1989) và Daubechies (1992) đã cùng công bố lý thuyết và triển khai ứng dụng phép biến đổi này.

Hàm wavelet được định nghĩa bởi (1), với a là tham số tỷ lệ và b là tham số dịch chuyển vị trí.

$$\Psi_{ab}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \tag{1}$$

Khi đó, phép biến đổi wavelet của tín hiệu f(t) được xác định là:

$$W(a, b) = \int_{-\infty}^{\infty} \Psi_{ab}(t) f^*(t) dt = \langle \Psi_{ab}, f \rangle \tag{2}$$

với $a \in R^+$, $b \in R$ và $f^*(t)$ là liên hiệp phức của f(t). Khi a, b là các tham số liên tục, thì (2) gọi là biến đổi wavelet liên tục. Ngược lại, nếu a, b được rời rạc hóa và được lấy mẫu như sau:

$$a = a_0^m; b = nb_0 a_0^m \tag{3}$$

thì (1) được biểu diễn thành:

$$\Psi_{mn}(t) = a_0^{-m/2} \Psi(a_0^{-m} t - b_0) \tag{4}$$

với $m, n \in Z$. Khi đó, $\{\Psi_{mn}\}$ là một thể wavelet khác và có thể dùng để biểu diễn tín hiệu f(t) bất kỳ dưới dạng sau (Akansu, 1994):

$$f(t) = \sum_m \sum_n d_{m,n} \Psi_{mn}(t) \tag{5}$$

trong đó, $d_{m,n}$ được gọi là hệ số wavelet, được xác định bởi:

$$d_{m,n} = \langle f(t), \Psi_{mn}(t) \rangle = \frac{1}{a_0^{m/2}} \int f(t) \Psi(a_0^{-m} t - nb_0) dt \tag{6}$$

Theo Akansu (1994), thay (6) vào (5), ta có:

$$f(t) = \sum_{n=-\infty}^{\infty} C_{L,n} a_0^{-L/2} \phi\left(\frac{t}{a_0^L} - nb_0\right) + \sum_{m=1}^L \sum_{n=-\infty}^{\infty} d_{m,n} a_0^{-m/2} \Psi\left(\frac{t}{a_0^m} - nb_0\right) \tag{7}$$

Với:

$$C_{L,n} = \langle f(t), \phi_{L,n}(t) \rangle = \frac{1}{a_0^{L/2}} \int f(t) \phi(a_0^{-L} t - nb_0) dt \tag{8}$$

Hàm $\psi(\cdot)$ và $\phi(\cdot)$ trong (7) cho phép biểu diễn tín hiệu f(t) dạng hai thành phần, đó là thành phần xấp xỉ A (Approximation), tương ứng với tín hiệu tần số thấp và thành phần chi tiết D (Detail) tương ứng với tín hiệu tần số cao. Với cách biểu diễn này, biến đổi wavelet được xem như tổng hợp của 2 bộ lọc, gồm: bộ lọc thông cao sử dụng hàm wavelet $\psi(\cdot)$ và bộ lọc thông thấp sử dụng hàm $\phi(\cdot)$. Mỗi quan hệ giữa hai hàm này như sau (Akansu, 1994):

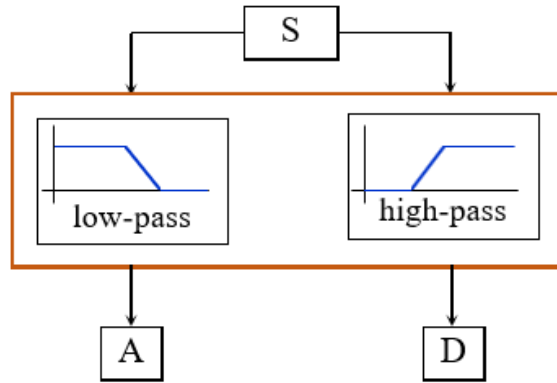
$$\phi(t) = 2 \sum_n h_0(n) \phi(2t - n) \tag{9}$$

$$\Psi(t) = 2 \sum_n h_1(n) \phi(2t - n) \tag{10}$$

trong đó $\{h_o(n)\}$ và $\{h_1(n)\}$ tương ứng là hệ số của bộ lọc thông thấp và bộ lọc thông cao.

Với cơ chế này, biến đổi wavelet trên tín hiệu nguồn S hoạt động như một bộ lọc ghép, gồm bộ lọc

hạ thông (low-pass filter) cho ra thành xấp xỉ A và bộ lọc thượng thông (high-pass filter) cho ra thành phần chi tiết D, như biểu diễn trên Hình 2.



Hình 2. Nguyên tắc lọc tín hiệu một cấp bằng wavelet

2.3. Chuẩn hóa biên độ tín hiệu

Tập dữ liệu ghi âm ở Bảng 1, được thực hiện lọc nhiễu bằng phép biến đổi wavelet. Tuy nhiên, biên độ tín hiệu ghi âm không đều nhau, do mỗi hiện tượng hư hỏng có tiếng ồn phát ra lớn, nhỏ khác nhau. Để đảm bảo biên độ tín hiệu giữa các tập tin âm thanh bằng nhau, tập dữ liệu âm thanh cần được chuẩn hóa biên độ về giá trị $[-1, 1]$.

Nếu tín hiệu âm thanh $f(t)$ có biên độ lớn nhất là f_{max} và biên độ nhỏ nhất là f_{min} , tức là $f(t) \in [f_{max}, f_{min}]$, thì tín hiệu chuẩn hóa $y(t)$ thu được từ $f(t)$ được thực hiện như sau (Chen, 2019):

$$y(t) = \begin{cases} f(t) \times f_{max}^{-1}, & \text{if } f_{max} > |f_{min}| \\ -f(t) \times f_{min}^{-1}, & \text{if } f_{max} \leq |f_{min}| \end{cases} \quad (11)$$

2.4. Tạo ảnh phổ tín hiệu bằng phép biến đổi wavelet

Việc tạo ảnh phổ tín hiệu âm thanh động cơ được xây dựng dựa trên nền tảng phép biến đổi wavelet liên tục, theo (2), nhưng triển khai trong miền tần số. Morse wavelet (Lilly et al., 2012) là phép biến đổi được sử dụng trong nghiên cứu này, với hàm wavelet thay vì sử dụng như (1) thì được định nghĩa như sau:

$$U(\omega) a_{\beta\gamma} \omega^\beta e^{-\omega\gamma} = \int_{-\infty}^{\infty} \Psi_{\beta\gamma}(t) e^{-i\omega t} dt = \quad (12)$$

với $a_{\beta\gamma}$ là hệ số chuẩn hóa, $U(\omega)$ là hàm nấc đơn vị, β và γ là 2 tham số thuộc tính định hình của hàm Morse wavelet. Khi tăng β và γ hàm wavelet (10)

dao động nhiều hơn trong miền thời gian và hẹp hơn trong miền tần số. Với đặc điểm này, việc dịch chuyển 2 tham số β và γ cho phép xác định được tần số đỉnh $\omega_{\beta\gamma peak}$ của tín hiệu và độ rộng hàm wavelet (Lilly et al., 2012):

$$\omega_{\beta\gamma peak} \equiv \left(\frac{\beta}{\gamma}\right)^{1/\gamma} \quad (13)$$

$$P_{\beta\gamma} = \sqrt{-\frac{\omega_{\beta\gamma peak}^2 \Psi_{\beta\gamma}''(\omega_{\beta\gamma peak})}{\Psi_{\beta\gamma}(\omega_{\beta\gamma peak})}} = \sqrt{\beta\gamma} \quad (14)$$

Trong đó, $\Psi_{\beta\gamma}''$ là đạo hàm cấp hai của $\Psi_{\beta\gamma}$ theo ω .

Gọi $W_f(a, b)$ là biến đổi Morse wavelet của tín hiệu âm thanh $f(t)$ với tham số tỷ lệ a và vị trí b . Phép biến đổi wavelet cho phép xác định mật độ năng lượng cục bộ của tín hiệu $f(t)$ tại mỗi cửa sổ wavelet (Mallat, 2009):

$$P_{W_f(a,\xi)} = |W_f(a, b)|^2 = \left|W_f\left(a, \frac{\omega_{\beta\gamma peak}}{\xi}\right)\right|^2 \quad (15)$$

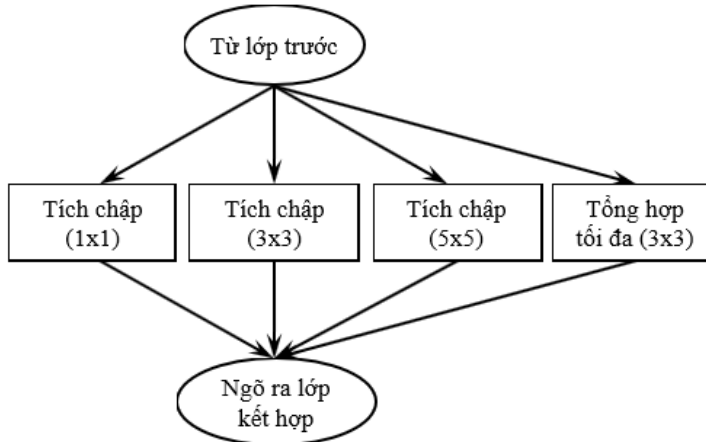
với $\xi = \frac{\omega_{\beta\gamma peak}}{b}$ là tâm mỗi cửa sổ wavelet. Biểu diễn mật độ năng lượng $P_{W_f(a,\xi)}$ theo thời gian khi trượt cửa sổ wavelet trên toàn tín hiệu $f(t)$ ta thu được ảnh phổ tín hiệu.

2.5. Mạng neuron tích chập GoogLeNet

GoogLeNet là một mạng neuron nhân tạo với 22 lớp học sâu, được tập đoàn Google kết hợp với nhiều viện, trường đề nghiên cứu và đề xuất vào năm 2014 (Szegedy et al., 2015; Mishra et al., 2020). Sức

mạnh của GoogLeNet là các dữ liệu đầu vào không cần thiết kể một bộ tiền xử lý ảnh như: lọc ảnh và trích xuất dữ liệu đặc trưng (Zhou et al., 2017). Kích thước của các ảnh huấn luyện và nhận dạng đều được thiết lập ở chuẩn $224 \times 224 \times 3$ điểm ảnh (pixels). Theo Arora et al. (2014), đặc điểm chính của mạng học sâu GoogLeNet gồm các lớp tích chập (convolution layer), các lớp tổng hợp tối đa (max pooling layer) và trọng tâm là các lớp kết hợp

(inception layer). Trong mỗi lớp kết hợp, kích thước của các lớp tích chập là cố định. Theo đó, lớp tích chập (1×1) , (3×3) , (5×5) và lớp tổng hợp tối đa (3×3) được thực hiện theo cách song song ở đầu vào và đầu ra của chúng, được xếp chồng lên nhau để tạo ra ngõ ra cuối cùng (Hình 3). Với phương pháp này, các bộ lọc tích chập có các kích thước khác nhau sẽ xử lý các đối tượng ở nhiều tỷ lệ, tốt hơn các phương pháp trước đó.



Hình 3. Cấu trúc của lớp kết hợp (Inception layer)

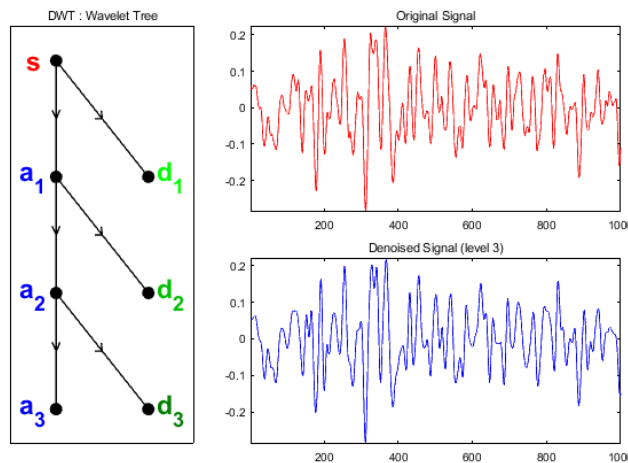
2.6. Khảo sát các thông số huấn luyện mạng

Các thông số tham gia quá trình huấn luyện mạng neuron tích chập có ảnh hưởng đến hiệu quả huấn luyện và khả năng nhận dạng của nó. Theo Mishra et al. (2020), hai tham số *miniBatchSize* và *MaxEpochs* có ảnh hưởng quan trọng đến thời gian huấn luyện và hiệu suất của mạng. Vì vậy, trong phần thực nghiệm, nghiên cứu này sẽ thay đổi các tham số trên, nhằm khảo sát và tìm kiếm bộ giá trị *miniBatchSize* và *MaxEpochs* khả dĩ nhất cho bài toán nhận diện sự cố động cơ điện.

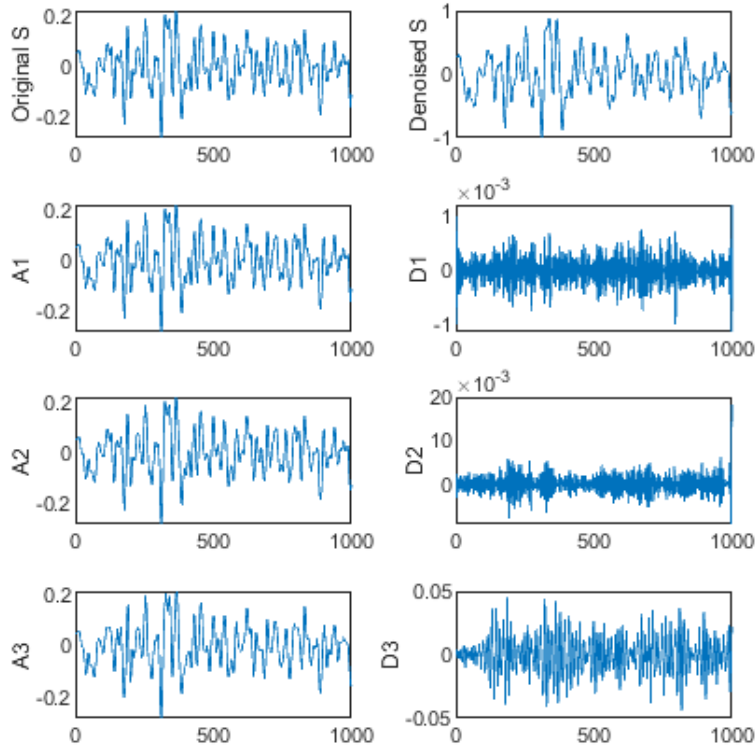
3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Tiền xử lý tín hiệu

Phần thực nghiệm của nghiên cứu này được triển khai trên phần mềm MATLAB. Công cụ *wavelet* (*wavelet toolbox*) của MATLAB hỗ trợ chức năng *waveletAnalyzer* cho phép minh họa cấu trúc cây của bộ lọc wavelet 3 bậc, như Hình 4. Thành phần tín hiệu xấp xỉ A và thành phần chi tiết D của tín hiệu âm thanh S theo cơ chế ở Hình 2 và các phép biến đổi (9), (10), được minh họa trên Hình 5.



Hình 4. Cấu trúc cây của bộ lọc 3 cấp bằng wavelet



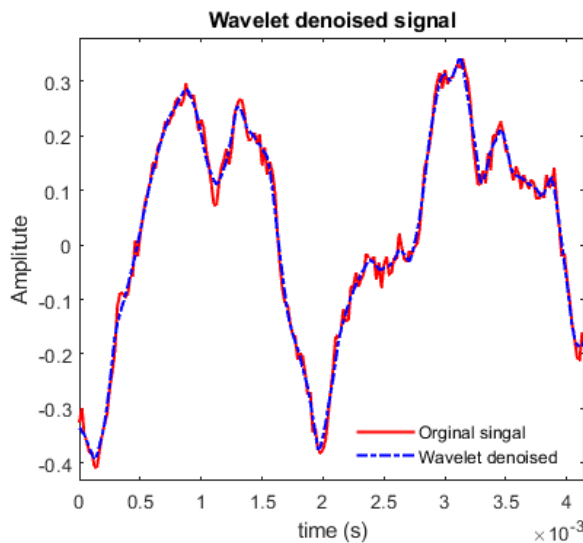
Hình 5. Thành phần xấp xỉ A và chi tiết D của tín hiệu âm thanh qua lọc wavelet

Nếu gọi s là tín hiệu âm thanh gốc, sd là tín hiệu âm thanh được lọc nhiễu, thì MATLAB hỗ trợ lệnh đơn giản để lọc nhiễu như sau:

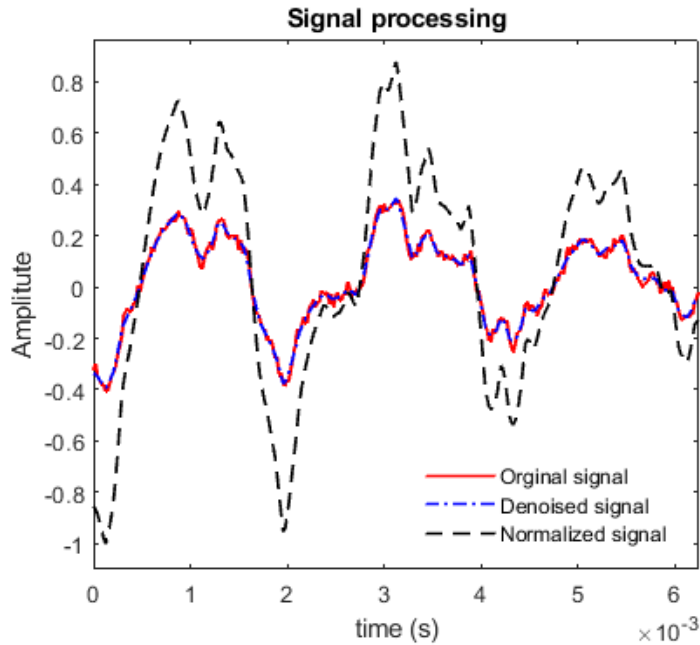
```
[s, fs] = audioread('<path><file>');
```

```
Sd = wdenoise(s,5,'DenoisingMethod','BlockJS');
```

Hình 6 minh họa kết quả lọc nhiễu bằng phép biến đổi wavelet. Hình 7 minh họa một đoạn tín hiệu âm thanh sau khi lọc nhiễu và chuẩn hóa về giá trị $[-1, 1]$. Tập dữ liệu âm thanh sau khi chuẩn hóa biên độ, sẽ được tiếp tục phân đoạn để tạo ảnh phổ bằng phép biến đổi wavelet.



Hình 6. Kết quả lọc nhiễu tín hiệu âm thanh động cơ bằng phép wavelet



Hình 7. Minh họa kết quả chuẩn hóa biên độ tín hiệu âm thanh

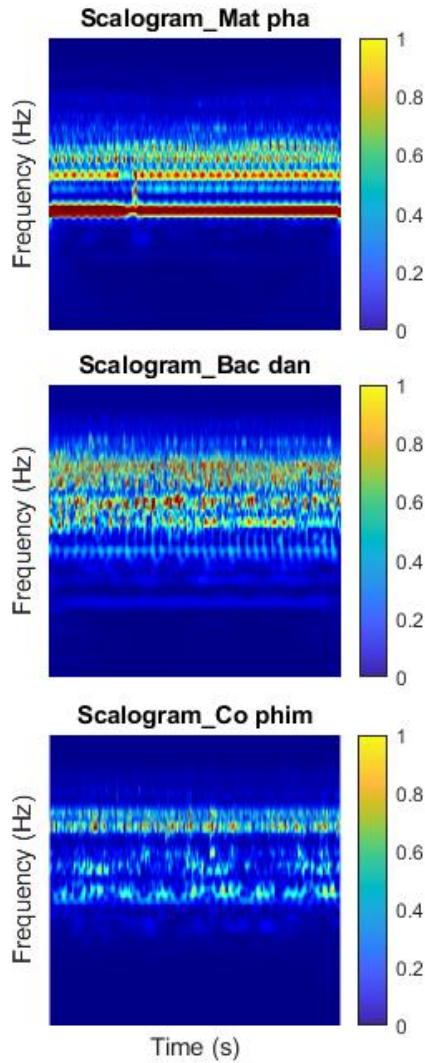
Công cụ wavelet của MATLAB cũng hỗ trợ mạnh mẽ việc dựng ảnh phổ tín hiệu bằng một số lệnh đơn giản. Hình 8 minh họa phổ ảnh tín hiệu âm thanh tương ứng với trường hợp động cơ mất pha, hồng bạc đạn và cọ phim cách điện. Mỗi tập tin dữ liệu âm thanh được cắt thành nhiều phân đoạn, mỗi phân đoạn có độ dài 65.536 mẫu (tần số lấy mẫu là 44.100 Hz). Mỗi phân đoạn tín hiệu này được sử dụng để tạo ra một ảnh phổ.

Dữ liệu âm thanh ở Bảng 1, tương ứng với 3 tình huống sự cố thông thường của động cơ được định nghĩa bằng các nhãn tương ứng là {'Mat_pha',

'Co_phim', 'Hong_bacdan'}, cũng như được tiền xử lý và chia ra thành các tập con ở Bảng 2. Trong đó, tập dữ liệu dùng để huấn luyện mạng, gọi là *Training_dataset* gồm {2.124, 2.203, 2.253} ảnh; tập dữ liệu dùng để kiểm tra mạng ngay trong quá trình huấn luyện (validation) gọi là *Validating_dataset*, gồm {531, 551, 564} ảnh; và tập dữ liệu độc lập dùng để kiểm thử gọi là *Testing_dataset* gồm {303, 391, 402} ảnh. Mạng neuron học sâu sẽ được huấn luyện bằng tập *Training_dataset* và *Validating_dataset*. Sau khi huấn luyện, mạng được đánh giá khả năng nhận dạng bằng tập *Testing_dataset*.

Bảng 2. Các tập dữ liệu ảnh phổ

Thông tin	Ký hiệu tập dữ liệu	Tình huống		
		Mất pha	Cọ phim	Hồng bạc đạn
Tập dữ liệu huấn luyện (ảnh)	Training_dataset	2.124	2.203	2.253
Tập dữ liệu kiểm tra khi huấn luyện (ảnh)	Validating_dataset	531	551	564
Tập dữ liệu độc lập để kiểm thử (ảnh)	Testing_dataset	303	391	402

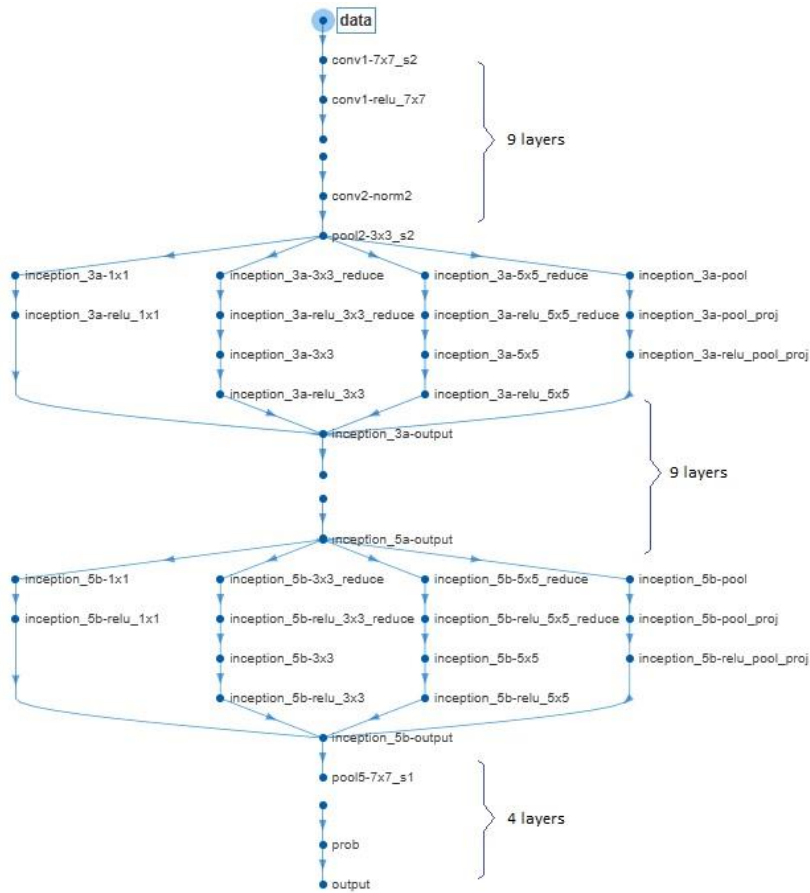


Hình 8. Minh họa ảnh phổ tín hiệu âm thanh động cơ

3.2. Huấn luyện mạng neuron học sâu

Mạng neuron học sâu được sử dụng trong thực nghiệm này là mạng GoogLeNet. MATLAB dàn trải 22 lớp học sâu của mạng GoogLeNet thành 144 lớp mềm với cấu trúc như Hình 9. Ngoài ra, hàm

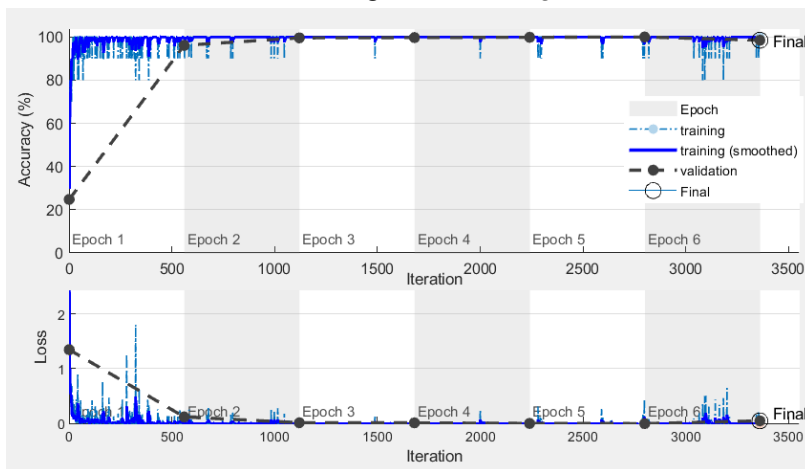
analyzeNetwork cho phép hiển thị cấu trúc mạng một cách trực quan đến từng lớp mềm. Ưu điểm của GoogLeNet là mạng này đã được huấn luyện trước với 10.000 ảnh thông dụng, do đó, khi tiếp tục huấn luyện với tập dữ liệu mới, mạng dễ dàng hội tụ (Mishra et al., 2020).



Hình 9. Cấu trúc của GoogLeNet trong MATLAB

Tập dữ liệu *Training_dataset* và *Validating_dataset* (Bảng 2) được sử dụng để huấn luyện và kiểm tra mạng neuron GoogLeNet. Cấu hình máy tính được sử dụng để huấn luyện mạng là: Dell Vostro 3671, Intel Core i7, CPU (8 nhân) 3.0GHz, 16 GB RAM, Windows 10-64 bit. Thông

số cơ bản được cài đặt để huấn luyện mạng gồm: *miniBatchSize* = 10 và *MaxEpochs* = 6. Tiến trình huấn luyện mạng được minh họa trên Hình 10, với thời gian huấn luyện trên tập *Training_dataset* là 54 phút 43 giây và tỷ lệ chính xác khi kiểm tra với tập *Validating_dataset* là 98,12%.



Hình 10. Huấn luyện mạng GoogLeNet với tập ảnh phổ âm thanh động cơ

3.3. Khảo sát các tham số huấn luyện mạng

Như đã đề cập, bộ tham số huấn luyện $\{miniBatchSize, MaxEpochs\}$ sẽ được thay đổi để đánh giá hiệu suất huấn luyện mạng. Ở loạt khảo sát đầu tiên, nghiên cứu này giữ nguyên giá trị mặc định $miniBatchSize = 10$ của MATLAB và tiến hành thay đổi tham số $MaxEpochs$ lần lượt các giá trị: 4, 6, 8, 10. Kết quả khảo sát thứ nhất được trình bày trong Bảng 3.

Kết quả ở Bảng 3 cho thấy tại tham số $MaxEpochs = 4$, thời gian huấn luyện mạng trên tập $Training_dataset$ là ngắn nhất (36 phút 28 giây) và tỷ lệ kiểm tra trên tập $Validating_dataset$ ngay trong quá trình huấn luyện đạt 98,63%. Như vậy, có thể thấy tham số $MaxEpochs = 4$ là giá trị tốt nhất trong khảo sát này.

Bảng 3. Khảo sát 1 – giữ nguyên $miniBatchSize = 10$ và thay đổi $MaxEpochs$

miniBatchSize = 10	MaxEpochs			
	4	6	8	10
Thời gian huấn luyện	36'28"	54'43"	72'3"	90'22"
Tỷ lệ chính xác khi kiểm tra trên tập Validating_dataset	98,63%	98,12%	98,31%	97,11%

Khảo sát thứ hai được tiến hành bằng cách giữ nguyên giá trị tham số $MaxEpochs = 4$ và tiến hành thay đổi tham số $miniBatchSize$ lần lượt các giá trị là 6, 8, 10, 12. Kết quả khảo sát thứ hai được trình bày trong Bảng 4.

tập $Training_dataset$ là 36 phút 28 giây, không chênh lệch đáng kể so với các trường hợp còn lại. Tuy nhiên, tỷ lệ kiểm tra trên tập $Validating_dataset$ ngay trong quá trình huấn luyện đạt giá trị cao nhất là 98,23%. Có thể thấy $miniBatchSize = 10$ là giá trị tốt nhất trong khảo sát này.

Kết quả ở Bảng 4 cho thấy tại tham số $miniBatchSize = 10$, thời gian huấn luyện mạng trên

Bảng 4. Khảo sát 2 – giữ nguyên $MaxEpochs$ và thay đổi $miniBatchSize$

MaxEpochs = 4	miniBatchSize			
	6	8	10	12
Thời gian huấn luyện	35'22"	35'13"	36'28"	37'06"
Tỷ lệ chính xác khi kiểm tra trên tập Validating_dataset	96,17%	97,26%	98,23%	96,79%

Kết quả ở Bảng 3 và 4 cho thấy cho thấy bộ tham số $\{miniBatchSize=10, MaxEpochs=4\}$ cho hiệu suất huấn luyện mạng ở bài toán này là tốt nhất. Với cấu trúc mạng này, nghiên cứu tiến hành đánh giá hiệu quả nhận dạng trên tập dữ liệu $Testing_dataset$ ở Bảng 2.

để nhận diện sự cố động cơ điện trên tập dữ liệu âm thanh độc lập $Testing_dataset$. Hình 11, 12 và 13 lần lượt minh họa kết quả nhận diện sự cố mất pha, cọ phim cách điện và hỏng bạc đạn, với tỷ lệ chính xác được trình bày ở Bảng 5. Mã QR-Code trên Hình 14 dẫn đến clip minh họa chương trình nhận diện sự cố động cơ điện bằng mạng GoogLeNet.

3.4. Đánh giá khả năng nhận diện sự cố động cơ của mạng

Mạng GoogLeNet đã huấn luyện được ứng dụng

Bảng 5. Đánh giá tỷ lệ chính xác trên tập dữ liệu độc lập

Kiểm thử trên tập $Testing_dataset$	Tình huống		
	Mất pha	Cọ phim	Hỏng bạc đạn
Số lượng (ảnh)	303	391	402
Tỷ lệ chính xác (%)	95,17	94,21	98,64

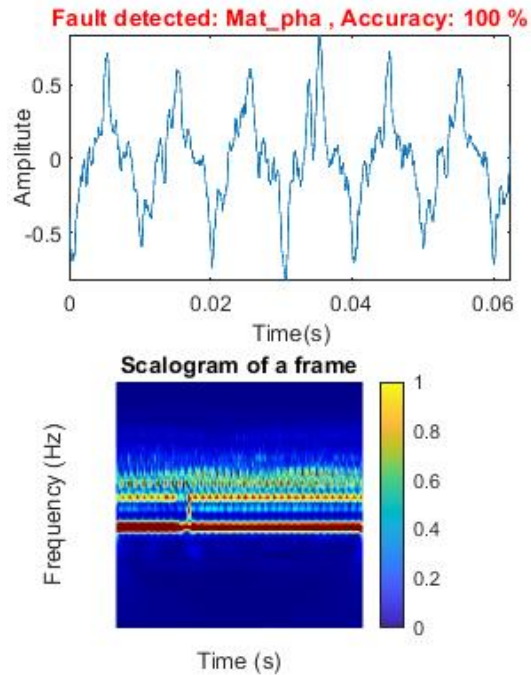
3.5. Thảo luận

Kết quả nhận diện đúng sự cố động cơ điện trên tập dữ liệu độc lập $Testing_dataset$ cho tỷ lệ chính xác dao động từ 94,21% đến 98,64%, tùy tình huống sự cố. Tuy nhiên, điểm hạn chế của nghiên cứu này

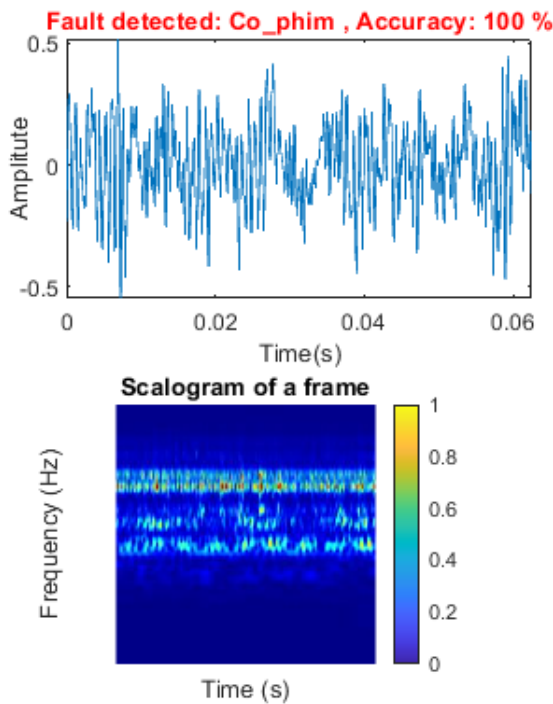
là chỉ mới thu thập dữ liệu sự cố trên một động cơ điện, với một vài loại hư hỏng thường gặp. Thời gian tới, nhóm nghiên cứu sẽ tiếp tục thu thập dữ liệu của nhiều loại động cơ điện khác nhau và bố trí ghi âm ngay trong môi trường công nghiệp. Ngoài ra, dữ liệu cũng sẽ được thu thập với các tốc độ động cơ

khác nhau và khoảng cách ghi âm khác nhau để đánh giá toàn diện hơn về giải pháp này. Với nền tảng lọc nhiễu bằng wavelet mạnh mẽ, nhóm nghiên cứu tin tưởng rằng có thể lọc được các tạp âm khác, trước khi nhận diện. Ngoài ra, khả năng xây dựng một ứng

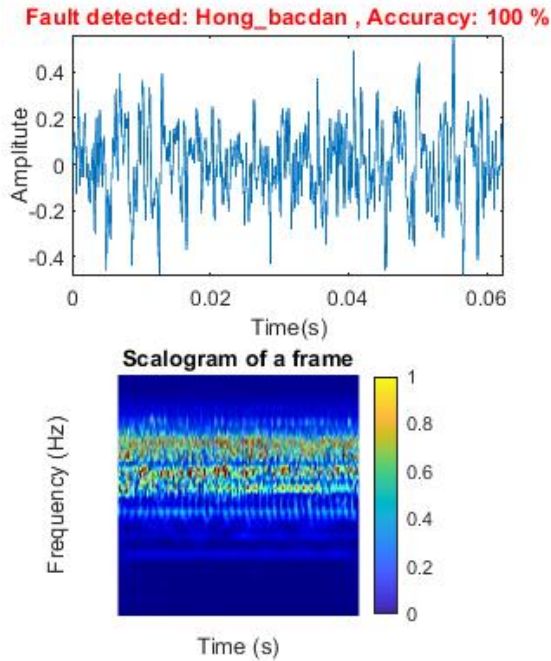
dụng cài đặt được trên điện thoại di động thông minh để hỗ trợ các doanh nghiệp chẩn đoán sớm sự cố động cơ điện là một hướng phát triển tiềm năng mà nhóm đang hướng đến.



Hình 11. Kết quả nhận diện sự cố mất pha



Hình 12. Kết quả nhận diện sự cố cộ phim cách điện



Hình 13. Kết quả nhận diện sự cố hồng bạc đạn



Hình 14. Clip minh họa kết quả nhận diện sự cố động cơ điện

4. KẾT LUẬN VÀ ĐỀ NGHỊ

Bài báo trình bày phương pháp chẩn đoán lỗi động cơ điện dựa trên nguyên tắc nhận diện ảnh phổ tín hiệu âm thanh thu được trong lúc vận hành dùng mạng neuron học sâu và kỹ thuật lọc nhiễu, tạo ảnh phổ 2 chiều bằng phép biến đổi wavelet. Việc bố trí thí nghiệm để ghi âm tiếng ồn phát ra trong lúc động cơ vận hành bằng điện thoại di động đã được thiết lập để ghi âm các tình huống sự cố thông dụng, bao gồm: vỡ bạc đạn, mất pha và cọ phim cách điện. Dữ liệu âm thanh sau đó được tiền xử lý thông qua việc lọc nhiễu bằng phép biến đổi wavelet và chuẩn hóa biên độ tín hiệu về giá trị [-1, 1]. Tập dữ liệu sau khâu tiền xử lý được biến đổi thành tập ảnh phổ tần số hai chiều, bằng phép biến đổi Morse wavelet. Tập

ảnh phổ được chia làm 2 phần, tương ứng 80% số ảnh dùng để xây dựng tập huấn luyện (*Training_dataset*) và 20% số ảnh dùng để xây dựng tập kiểm tra (*Validating_dataset*). Ngoài ra, một số tập tin âm thanh gốc độc lập cũng được chuẩn bị để kiểm tra khả năng nhận diện lỗi của giải thuật (*Testing_dataset*). Mạng neuron tích chập GoogLeNet được sử dụng và khảo sát hiệu quả huấn luyện thông qua thay đổi bộ tham số cơ bản {*miniBatchSize*, *MaxEpochs*}, nhằm xác định hiệu quả huấn luyện tốt nhất. Mạng GoogLeNet sau khi huấn luyện với tham số {*miniBatchSize* = 10, *MaxEpochs* = 4} được kiểm tra trên tập dữ liệu âm thanh độc lập. Kết quả cho thấy mạng có khả năng nhận diện 3 sự cố động cơ đã thiết lập đạt tỷ lệ chính xác 94,21%.

Trong thời gian tới, nghiên cứu sẽ được tiếp tục phát triển trên tập dữ liệu của nhiều loại động cơ điện khác nhau và bố trí ghi âm ngay trong môi trường công nghiệp ở tốc độ động cơ và khoảng cách thu âm khác nhau. Với nền tảng tiền xử lý tín hiệu mạnh mẽ bằng wavelet, nhóm nghiên cứu tin tưởng có thể lọc được các tạp âm khác. Đồng thời, khả năng xây dựng một ứng dụng cài đặt được trên điện thoại di động thông minh để hỗ trợ các doanh nghiệp chẩn đoán sớm sự cố động cơ điện là một hướng đi tiềm năng mà nhóm đang hướng đến.

TÀI LIỆU THAM KHẢO

- Akansu, A. N. (1994). Wavelets and filter banks A signal processing perspective. *IEEE Circuits and Devices Magazine*, 10, 14-18.
- Arora, S., Bhaskara, A., Ge, R., & Ma, T. (2014). Provable bounds for learning some deep representations. *Proc. of the 31st Inter. Confer. on Machine Learning, in PMLR*, 32(1), 584-592.
- Bonnett, A.H., & Soukup, G.C. (1992). Cause and analysis of stator and rotor failures in three-phase squirrel-cage induction motors. *IEEE Transactions on Industry Applications*, 28(4), 921-937. DOI: 10.1109/28.148460.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia, SIAM, 1st Ed. ISBN: 978-0-89871-274-2.
- Chen, Y.-W. (2021). Audio Normalization. *MATLAB Central File Exchange* (retrieved 10/8/2021).
- Ciszewski, T., Gelman, L., & Swedrowski, L. (2016). Current-based higher-order spectral covariance as a bearing diagnostic feature for induction motors. *Insight - Non-Destructive Testing and Condition Monitoring*, 58(8), 431-434.
- Glowacz, A., & Glowacz, Z. (2017). Diagnosis of the three-phase induction motor using thermal imaging. *Infrared Physics & Technology*, 81, 7-16, <https://doi.org/10.1016/j.infrared.2016.12.003>.
- Glowacz, A., Glowacz, W., Glowacz, Z., Kozik, J., Gutten, M., Korenciak, D., Khan, Z.F., Irfan, M., & Carletti, E. (2017). Fault diagnosis of three phase induction motor using current signal, MSAF-Ratio15 and selected classifiers. *Arch Metall Mater*, 62(4), 2413-2419.
- Glowacz, A., Glowacz, W., Glowacz, Z., & Kozik, J. (2018). Early fault diagnosis of bearing and stator faults of the single-phase induction motor using acoustic signals. *Measurement*, 113, 1-9. <http://dx.doi.org/10.1016/j.measurement.2017.08.036>.
- Guo, P., Infield, D., & Yang, X. (2012). Wind turbine generator condition monitoring using temperature trend analysis. *IEEE Trans. Sustain. Energy*, 3(1), 124-133.
- Haar, A. (2010). Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen*, 69(3), 331-371. DOI:10.1007/BF01456326.
- Henao, H., Capolino, G.A., Fernandez-Cabanas, M., Filippetti, F., Bruzzese, C., Strangas, E., Pusca, R., Estima, I., Riera-Guasp, M., & Hedayati-Kia, S. (2014). Trends in Fault Diagnosis for Electrical Machines A Review of Diagnostic Techniques. *IEEE Industrial Electronics Magazine*, 8(2), 31-42. DOI:10.1109/MIE.2013.2287651.
- Ince, T., Kiranyaz, S., Eren, L., Askar, M., & Gabbouj, M. (2016). Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks. *IEEE Transactions on Industrial Electronics*, 63(11), 7067-7075. DOI:10.1109/TIE.2016.2582729.
- Kim, D., Kim, H., Hong, J., & Park, C. (2014). Estimation of Acoustic Noise and Vibration in an Induction Machine Considering Rotor Eccentricity. *IEEE Transactions on Magnetics*, 50(2), 857-860. DOI:10.1109/TMAG.2013.2285391.
- Kumar, A., Gandhi, C.P., Zhou, Y., Kumar, R., & Xiang, J. (2020). Improved deep convolution neural network (CNN) for the identification of defects in the centrifugal pump using acoustic images. *Applied Acoustics*, 167, 107399. ISSN 0003-682X. <https://doi.org/10.1016/j.apacoust.2020.107399>.
- Lilly, J.M., & Olhede, S. C. (2020). Generalized Morse Wavelets as a Superfamily of Analytic Wavelets. *IEEE Transactions on Signal Processing*, 60(11), 6036-6041. DOI:10.1109/TSP.2012.2210890.
- Li, Y., Chai, F., Song, Z., & Li, Z. (2017). Analysis of Vibrations in Interior Permanent Magnet Synchronous Motors Considering Air-Gap Deformation. *Energies*, 10(9), 1259. <https://doi.org/10.3390/en10091259>.
- Mallat, S. G. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7), 674-693.
- Mishra, A., Ranjan, P., & Ujlayan, A. (2020). Empirical analysis of deep learning networks for affective video tagging. *Multimedia Tools Applications*, 79, 18611-18626, Springer. <https://doi.org/10.1007/s11042-020-08714-y>
- Qi, Y., Shen, C., Wang, D., Shi, J., Jiang, X., & Zhu, Z. (2017). Stacked Sparse Auto encoder-Based Deep Network for Fault Diagnosis of Rotating Machinery. *IEEE Access*, 5, 15066-15079.
- Sangeetha, P., & Hemamalini, S. (2017). Dyadic wavelet transform-based acoustic signal analysis for torque prediction of a three-phase induction motor. *IET Signal Processing*, 11(5), 604-612. DOI: 10.1049/iet-spr.2016.0165.
- Singh, S.K., Kumar, S., & Dwivedi, J.P. (2017). Compound fault prediction of rolling bearing using multimedia data. *Multimedia Tools Applications*, 76, 18771-18788. <https://doi.org/10.1007/s11042-017-4419-1>.
- Szegedy, C., Wei, L., Yangqing, J., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going Deeper with

- Convolutions. *IEEE Confer. on Computer Vision and Pattern Recognition – CVPR 2015* (pp. 1-9). DOI: 10.1109/CVPR.2015.7298594.
- Wen, L., Gao, L., & Li, X. (2017). A New Deep Transfer Learning Based on Sparse Auto-Encoder for Fault Diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(1), 136-144. DOI:10.1109/TSMC.2017.2754287.
- Xia, M., Li, T., Xu, L., Liu, L., & de Silva, C.W. (2018). Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks. *IEEE/ASME Transactions on Mechatronics*, 23(1), 101-110. DOI:10.1109/TMECH.2017.2728371.
- Yang, T., Pen, H., Wang, Z., & Chang, C.S. (2016). Feature Knowledge Based Fault Detection of Induction Motors Through the Analysis of Stator Current Data. *IEEE Transactions on Instrumentation and Measurement*, 65(3), 549-558. DOI: 10.1109/TIM.2015.2498978.
- Zhou, B., Khosla, A., Lapedriza, A., Torralba, A., & Oliva, A. (2014). Places: An image database for deep scene understanding. *Journal of Vision*, 17(10), 296, 1-12. DOI: <https://doi.org/10.1167/17.10.296>.