

Report

Sinh viên	Đỗ Thành Vinh
MSSV	20110112
Tên đồ án	Chẩn đoán lỗi ổ lăn từ tín hiệu rung dựa trên kết hợp đặc trưng sâu và nông
Dataset	CWRU Bearing Dataset (Case Western Reserve University)
Bài báo tham chiếu	<i>Vibration signal analysis for rolling bearings faults diagnosis based on deep-shallow features fusion</i>
Link bài báo	nature.com/articles/s41598-025-93133-y
Link dataset	engineering.case.edu/bearingdatacenter/download-data-file
Tài liệu môn học	Bishop — <i>Pattern Recognition and Machine Learning</i> (2006)

1) Abstract

Đồ án giải bài toán phân loại tình trạng ổ lăn từ tín hiệu rung. Đầu vào là tín hiệu rung 1 chiều $x[n]$ (drive-end, $f_s = 48 \text{ kHz}$), đầu ra là nhãn y thuộc 10 lớp, đánh giá bằng accuracy, F1-macro và confusion matrix. Em chuyển tín hiệu sang ảnh phổ log-mel (96×64), huấn luyện mô hình CNN và thử kết hợp đặc trưng sâu (CNN) với đặc trưng nông (MBH-LPQ) bằng fusion theo trọng số. Ngoài thực nghiệm cùng tải, em mở rộng kiểm tra domain shift theo tải (load 0–3) bằng thiết lập train load A, test load B.

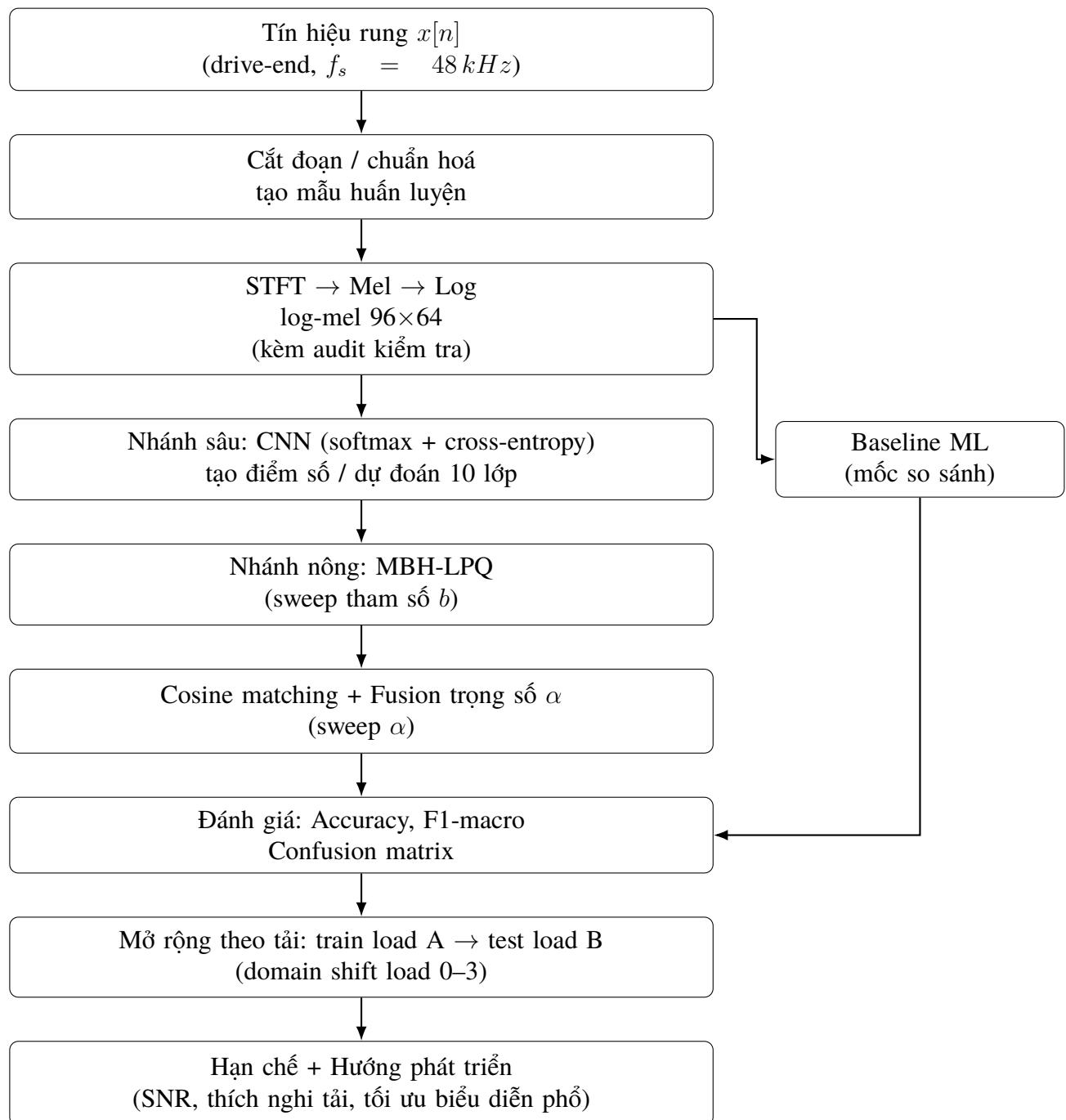
2) Ý nghĩa ứng dụng và đóng góp của hướng tiếp cận (bài báo)

Ý nghĩa ứng dụng: Chẩn đoán rung phục vụ bảo trì dự đoán: phát hiện sớm hư hỏng ổ lăn để giảm dừng máy và giảm rủi ro sự cố dây chuyền. Hướng làm này phù hợp bối cảnh công nghiệp vì chỉ cần tín hiệu rung (dễ thu thập) và cho đầu ra là nhãn trạng thái.

Đóng góp về mặt phương pháp (theo tinh thần bài báo):

- Đưa ra một pipeline rõ ràng: rung 1D → biểu diễn phổ (log-mel) → trích đặc trưng → phân loại.
- Kết hợp hai nguồn thông tin bổ sung: đặc trưng sâu (học mức cao từ CNN) và đặc trưng nông (texture cục bộ như LPQ/MBH-LPQ).
- Fusion theo trọng số giúp điều chỉnh mức đóng góp từng nhánh; đồng thời có thể phân tích/giải thích bằng confusion matrix và các sweep tham số (thay vì chỉ đưa một con số).

3) Trục xương sống (Sơ đồ)



4) Nội dung đồ án

Dữ liệu và thiết lập:

- Dataset CWRU, kênh drive-end, $f_s = 48kHz$, bài toán 10 lớp; thực nghiệm theo tải load 0–3.
- Chia train/test 80/20 (seed cố định). Có đánh giá cùng tải ($A=A$) và đánh giá chéo tải ($A \neq B$).

Biểu diễn đặc trưng:

- Chuyển $x[n]$ sang ảnh phổ log-mel 96×64 .
- Có output “audit log-mel” để kiểm tra sớm lỗi STFT/mel/chuẩn hoá trước khi chạy hàng loạt.

Mô hình và so sánh:

- Baseline ML: mốc so sánh theo các dạng phân lớp cơ bản trong môn.
- CNN: huấn luyện đa lớp (softmax + cross-entropy), đánh giá bằng confusion matrix.
- Fusion: fusion mức điểm (score-level) theo trọng số α , và đổi chiều thêm fusion mức đặc trưng (feature-level).

Theo bài báo (paper-core):

- Dùng cosine similarity để so khớp đặc trưng/điểm số.
- Quét α để chọn trọng số fusion.
- Quét tham số MBH-LPQ theo số block b ; kích thước đặc trưng tăng theo $\dim(f) = 256 \cdot b$.

Cải tiến và mở rộng:

- Cải tiến (ablation): thêm mốc nhánh nồng đơn giản để đo riêng phần texture.
- Mở rộng: domain shift theo tải bằng train load A, test load B (heatmap + confusion matrix minh họa).

5) Tóm tắt từng slide (Slide 2 → Slide 21)

1. **Slide 2 — Nội dung trình bày:** Liệt kê các phần: dữ liệu CWRU theo tải 0–3, bài báo tham chiếu, log-mel, baseline ML, CNN, fusion, sweep tham số theo bài báo, cải tiến/ablation, domain shift, hạn chế và hướng phát triển.
2. **Slide 3 — Bối cảnh, mục tiêu, đầu vào/dầu ra:** Nêu lý do (ô lăn hỏng gây dừng máy), mục tiêu theo trực hiểu → tái hiện → giải thích → cải tiến → mở rộng theo tải. Input $x[n]$ (drive-end, 48kHz), output nhãn 10 lớp; metric accuracy, F1-macro, confusion matrix; biểu diễn log-mel 96×64 .

3. **Slide 4 — Dữ liệu CWRU và thiết lập theo tải:** Dữ liệu .mat; thiết lập 48kHz, 10 lớp, train/test 80/20 (seed cố định). Thực nghiệm theo load 0–3 và đánh giá chéo tải; tóm tắt nhãm 10 lớp: Normal (1), IR (3), Ball (3), OR (3).
4. **Slide 5 — Output tiền xử lý: kiểm tra log-mel:** Mục tiêu audit: log-mel sinh đúng dải tần/cấu trúc trước khi huấn luyện toàn bộ, để phát hiện sớm lỗi STFT/mel/chuẩn hoá. Có công thức $S(t, m) = \log(\varepsilon + \sum_{\omega} |X(t, \omega)|^2 H_m(\omega))$.
5. **Slide 6 — Output log-mel theo tải: 2 ví dụ đổi chiều:** Đổi chiều train/test cùng tải (0→0) và khi điều kiện thay đổi (0→3). Nhận xét: đổi tải làm phân bố năng lượng theo dải tần thay đổi, là nguyên nhân trực quan của domain shift.
6. **Slide 7 — Output tổng hợp theo tải (load 0–3):** Bảng tổng hợp theo điều kiện train_load = test_load gồm: load, b, α, acc_test, f1_test. Kết quả: load 0/1/2 đạt 1.0000; load 3 đạt 0.9983. Cấu hình tốt nhất theo từng load: Fusion với $b = 1, \alpha = 0.1$ (trích từ results/phase10/domain_shift_summary.csv).
7. **Slide 8 — Output baseline ML (mốc so sánh):** Baseline là mốc trước khi chuyển sang CNN và fusion; mô hình baseline dùng các dạng phân lớp cơ bản trong môn. Có công thức accuracy: $\text{Acc} = \frac{1}{N} \sum_{i=1}^N \mathbf{1}[\hat{y}_i = y_i]$.
8. **Slide 9 — Output mô hình CNN:** Đánh giá bằng confusion matrix trên tập test; huấn luyện đa lớp dùng softmax và cross-entropy: $p(y = k|x) = \frac{e^{a_k}}{\sum_j e^{a_j}}, L = -\sum_i \log p(y_i|x_i)$.
9. **Slide 10 — Output fusion mức điểm (score-level):** Đánh giá bằng confusion matrix; fusion mức điểm trộn hai nhánh theo trọng số α : $s = \alpha s_{\text{sâu}} + (1 - \alpha)s_{\text{nông}}, 0 \leq \alpha \leq 1$.
10. **Slide 11 — Output so sánh hai kiểu fusion:** Đặt cạnh nhau score-level và feature-level (hai confusion matrix) để nhìn khác biệt khi gộp theo điểm số so với gộp theo vector đặc trưng.
11. **Slide 12 — Theo bài báo: cấu hình tốt nhất và quét α :** Trình bày confusion matrix của cấu hình theo bài báo và đồ thị quét α để chọn fusion. Nhắc công thức cosine similarity $\cos(u, u') = \frac{u^\top u'}{\|u\|\|u'\|}$ và công thức score fusion.
12. **Slide 13 — Theo bài báo: quét tham số MBH-LPQ:** Đồ thị độ chính xác thay đổi theo số block b ; tăng b giữ chi tiết cục bộ nhiều hơn nhưng tăng chiều: $\dim(f) = 256 \cdot b$.
13. **Slide 14 — Bảng số liệu đánh giá (accuracy và F1-macro):**
Bảng tóm tắt: CNN đạt Accuracy(test)=100.00%, F1-macro(test)=1.000; Fusion (score-level) đạt Accuracy(test)=100.00%, F1-macro(test)=1.000 (trích trực tiếp từ output trong thư mục results).
14. **Slide 15 — Kết quả fusion (nông + sâu):** Minh họa confusion matrix cho score-level và feature-level; giải thích ý nghĩa: score-level trộn điểm sau so khớp, feature-level nối vector rồi phân loại. Nhắc lại: $s = \alpha s_{\text{sâu}} + (1 - \alpha)s_{\text{nông}}$.
15. **Slide 16 — Output cải tiến: thêm mốc nhánh nông:** Cải tiến thử nghiệm: tạo thêm mốc nhánh nông đơn giản để đo riêng phần texture; output là confusion matrix của mốc này.
16. **Slide 17 — Output mở rộng theo tải: domain shift (tổng quan):** Heatmap: mỗi ô là kết quả train load A, test load B. Khi A khác B, độ chính xác thường giảm do khác phân phối theo tải.
17. **Slide 18 — Output domain shift: hai ví dụ confusion matrix:** So sánh train 0→test 0 với train 0→test 3. Khi đổi tải, nhầm lẫn tăng và confusion matrix loang hơn.

18. **Slide 19 — Mở rộng domain shift theo tải (0,1,2,3):** Trình bày ma trận domain shift (train load A→test load B) và minh họa một cặp cụ thể (train 0→test 3) bằng confusion matrix.
19. **Slide 20 — Hạn chế (có minh chứng):** Khi train/test khác tải, mức nhầm lẫn tăng rõ trên confusion matrix. Một số lớp gần nhau vẫn khó tách khi phổ tương tự.
20. **Slide 21 — Hướng phát triển (có minh chứng):** (i) Tối ưu biểu diễn phổ theo tải: thử cấu hình STFT, mel, chuẩn hoá và đổi chiều bằng output. (ii) Kiểm tra độ bền với nhiễu theo SNR và báo cáo bằng bảng số liệu. (iii) Xử lý domain shift: chọn tham số theo tải hoặc học thích nghi theo tải. Công thức: $\text{SNR}(\text{dB}) = 10 \log \left(\frac{P_s}{P_n} \right)$.