



تشخیص خرابی یاتاقان شعاعی شفت دوار با استفاده از طبقه‌بندی توسط طیف‌نگار صوتی و شبکه عصبی پیچشی

مهدی اوستاد، امیر راستی*، آرمین رزائی

گروه مهندسی مکانیک، دانشگاه تربیت مدرس، تهران، ایران

* ایمیل نویسنده مسئول: a.rasti@modares.ac.ir

اطلاعات مقاله

مقاله پژوهشی

دریافت: ۱۵ اردیبهشت ۱۴۰۴

پذیرش: ۱۹ مرداد ۱۴۰۴

چکیده

خرابی یاتاقان‌ها یکی از عوامل اصلی ازکارافتادگی در تجهیزات دوار محسوب می‌شود. از این رو، نیاز به روش‌های مؤثر برای شناسایی به موقع عیوب یاتاقان‌ها با هدف جلوگیری از خرابی‌های ناگهانی ماشین‌آلات، به طور فزاینده‌ای در حال افزایش است. مطالعات پیشین، روش‌هایی همچون اندازه‌گیری دما و پایش ارتعاشات را برای تشخیص خرابی مورد بررسی قرار داده‌اند، اما این رویکردها به دلیل حساسیت بالا به نویز، با محدودیت‌هایی مواجه‌اند. در نتیجه، پژوهشگران به بهره‌گیری از تحلیل سیگنال‌های صوتی روی آورده‌اند. در میان روش‌های نوین، ضرایب کپسترال فرکانس مل، طیف‌نگارهای صوتی، و شبکه‌های عصبی پیچشی دویعدی، توجه زیادی را به خود جلب کرده‌اند. با این حال، روش‌های مبتنی بر ضرایب کپسترال فرکانس مل به نرخ نمونه‌برداری بالا و پهنای باند فرکانسی گسترده‌ای نیاز دارند. در این پژوهش، روشی مبتنی بر طیف‌نگاری پیشنهاد شده است که تمرکز بر استخراج ویژگی‌های فرکانسی و کاهش نویز از طریق فیلترگذاری و مدل‌سازی خرابی توسط یک شبکه عصبی پیچشی بهینه‌شده متصل دارد. در این روش، ابتدا طیف‌نگاری با تمرکز بر یک باند فرکانسی باریک، تصاویر با وضوح پایین تولید می‌کند. این تصاویر توسط یک شبکه عصبی پیچشی شامل لایه‌های کانولوشنی و لایه‌های کاملاً متصل پردازش می‌شوند. نتایج آزمایش‌ها نشان می‌دهند که سیستم پیشنهادی با دقت ۹۹/۸۸٪ و پیچیدگی محاسباتی کاهش یافته، عملکرد قابل توجهی دارد. شبکه عصبی بهینه‌شده دارای ۶۲۲/۷۷ کیلوبایت پارامتر و 1.52×10^6 عملیات ممیز شناور در ثانیه است. این سامانه تشخیص خرابی حتی در شرایط متغیر سرعت دوران و نرخ نمونه‌برداری پایین نیز عملکرد مؤثری از خود نشان می‌دهد.

کلیدواژگان:

عیب‌یابی

طبقه‌بندی

یادگیری ماشین

شبکه عصبی پیچشی

یاتاقان شعاعی

Fault diagnosis radial bearing mounted on shaft using classification with audio spectrogram and convolutional neural network

Mehdi Ustad, Amir Rasti*, Armin Rezaei

Department of Mechanical Engineering, Tarbiat Modares University, Tehran, Iran

* Corresponding Author's Email: a.rasti@modares.ac.ir

Article Information

Original Research Paper

Received: 5 May 2025

Accepted: 10 August 2025

Keywords:

Fault Diagnosis

Classification

Machine Learning

Convolutional Neural Network

Radial Bearing

Abstract

Bearing failures are a primary cause of breakdowns in rotating machinery. As such, there is an increasing demand for effective bearing fault detection methods to prevent machinery failures. Previous studies have explored temperature measurement and vibration monitoring for fault detection, but these approaches face limitations due to noise interference. Consequently, researchers have turned to sound signal monitoring. Among modern techniques, Mel Frequency Cepstral Coefficients (MFCCs), audio spectrograms, and two-dimensional Convolutional Neural Networks (CNNs) have attracted significant interest. However, existing MFCC methods require high sampling rates and wide frequency bands. In this study, we propose a spectrogram-based method emphasizing frequency features and noise reduction using filters, integrated with an optimized CNN. The spectrogram analyzes a narrow frequency band and generates low-resolution images, which are then processed by a CNN designed with convolutional and fully connected layers. Experimental results demonstrate that the proposed system achieves 99.88% accuracy with reduced complexity. The optimized CNN has 622.77 kB of parameters and 1.53×10^6 FLOPs. This fault diagnosis system proves effective even under varying rotational frequencies and low sampling rates.

Please cite this article using:

Ustad M, Rasti A, Rezaei A. Fault diagnosis radial bearing mounted on shaft using classification with audio spectrogram and convolutional neural network. Iranian Journal of Manufacturing Engineering. 2025 Jul 23;12(5):1-10. doi: 10.22034/ijme.2025.521365.2077 [In Persian]

برای ارجاع به این مقاله از عبارت ذیل استفاده نمایید:

۱- مقدمه

یاتاقان‌ها وظیفه کاهش اصطکاک چرخشی و حمل بارهای شعاعی و محوری را در ماشین‌آلات دوار بر عهده دارند. این تجهیزات از جمله قطعاتی هستند که بیشترین خرابی را در ماشین‌آلات تجربه می‌کنند [۱، ۲]. در برخی موارد بویژه در ماشین‌های ابزار و دستگاه‌های خطوط تولید پیوسته مانند خطوط نورد و شکل‌دهی غلتکی، این خرابی می‌تواند باعث خرابی در قطعات تولیدی و هزینه سرسام‌آور گردد [۳-۶]. بنابراین، خرابی یاتاقان‌ها می‌تواند باعث آسیب جدی به ماشین‌آلات دوار شود و منجر به خسارات اقتصادی و افزایش خطرات ایمنی گردد [۷]. در چند دهه گذشته، مطالعات زیادی در زمینه تشخیص خرابی یاتاقان‌ها انجام شده است. روش‌های تشخیص خرابی یاتاقان‌ها را می‌توان به اندازه‌گیری دما، پایش صدا و پایش ارتعاش دسته‌بندی کرد [۸، ۹]. برای مثال اندازه‌گیری دما از تصاویر دوبعدی یاتاقان‌ها که با استفاده از دوربین تصویربرداری حرارتی مادون‌قرمز به دست می‌آید، برای تشخیص خرابی یاتاقان‌ها استفاده می‌شود [۱۰]. این درحالی است که پایش صدا از سیگنال‌های صوتی یاتاقان که با استفاده از میکروفون به ضبط می‌شود، برای تشخیص خرابی‌ها استفاده می‌کند [۱۱]. حسگرهای مورد استفاده در روش‌های متداول محدودیت‌هایی دارند. در مورد دوربین‌های تصویربرداری حرارتی، گردوغبار و ذرات هوا می‌توانند با اطلاعات تصویر تداخل کرده و داده‌برداری را با مشکل مواجه کنند. در میکروفون‌ها نیز، نویز خارجی سیگنال را تحریف می‌کند و صدا در فاصله اندازه‌گیری دورتر تضعیف می‌شود. ارتعاش‌سنج‌ها در مقابله با عوامل خارجی مقاوم هستند، زیرا آن‌ها سیگنال‌های ارتعاشی را در تماس با ماشین‌آلات دوار اندازه‌گیری می‌کنند. با این حال، این تجهیزات نیاز به نصب مستقیم دارند. همچنین برای پایش هر یاتاقان به یک ارتعاش‌سنج نیاز است [۱۲]. پایش صدا با پیش‌پردازش صوت و فیلترینگ نویز این محدودیت‌ها را تا حد زیادی مرتفع کرده است و هزینه اجرای این روش برای تحلیل خرابی یاتاقان‌ها با روش‌های ارتعاشی قابل‌مقایسه نیست [۱۳].

شبکه‌های عصبی پیچشی^۱ (CNN) دارای قابلیت‌های استخراج ویژگی‌های^۲ غیرخطی قدرتمندی هستند [۱۴]، بنابراین بسیاری از محققان اخیراً از CNN برای طبقه‌بندی سیگنال‌های صوتی یاتاقان استفاده کرده‌اند [۱۵-۱۷]. سیگنال‌های صوتی یاتاقان شامل اطلاعات زمانی و فرکانسی از جمله، تبدیل فوریه کوتاه‌مدت^۳ (STFT)، تبدیل موجک^۴ (WT) و طیف‌نگاری صوتی^۵ برای تحلیل سیگنال‌های صوتی از نظر زمان و فرکانس استفاده می‌شوند [۱۸]. در این میان، طیف‌نگارها یک روش ساده برای تبدیل سیگنال‌های صوتی به زمان و فرکانس است. بنابراین، طیف‌نگاری برای شبکه CNN کارآمدتر از WT و STFT است. شان و همکاران [۱۳] یک Mel-CNN را پیشنهاد کردند که یک تصویر MFCC^۶ و یک CNN را با استفاده از سیگنال‌های صوتی ترکیب می‌کند. در این پژوهش طیف‌نگاری مل به طور مؤثری ویژگی‌های خرابی یاتاقان را استخراج کرده است، درحالی‌که به دلیل همبستگی کم بین بردارها، و مدل CNN برای طبقه‌بندی تصویر به‌منظور دستیابی به طبقه‌بندی مؤثر خرابی‌ها ترکیب شد. با این حال، سیگنال صوتی ویژگی‌های خرابی یاتاقان و همچنین مقدار زیادی نویز سفید^۷ را شامل می‌شد. علاوه بر این، نویز در کارخانه‌های صنعتی به‌شدت زیاد است و طبقه‌بندی خرابی یاتاقان‌ها را با مشکل مواجه می‌کند. شیا و همکاران [۱۹] یک مدل CNN برای تشخیص خرابی در پایش ارتعاش پیشنهاد کردند. این تحقیق از روشی استفاده می‌کند که سیگنال‌های ارتعاشی چندکاناله جمع‌آوری شده را در یک ردیف قرار داده و اطلاعات زمانی و مکانی سیگنال‌های اصلی را حفظ می‌کند، به طوری که با اطلاعات ورودی به طور مؤثری خرابی را تشخیص می‌دهد. چالش اصلی، CNN‌های موجود در کاربرد تشخیص خرابی این است که آن‌ها به تعداد زیادی نمونه آموزشی برای دستیابی به دقت تشخیصی بالا نیاز دارند. در یک مطالعه، یک روش تشخیص خرابی یاتاقان بر اساس ضرایب کپسترال^۸ فرکانس مل (MFCCs) و یک CNN پیشنهاد شد [۲۰]. ابتدا، یک MFCC برای استخراج ویژگی‌های سیگنال خرابی معرفی شد. سپس، یک CNN با بلوک‌های پیچشی تغییرپذیر برای استخراج اطلاعات مکانی بین دسته‌ها و بهبود ویژگی‌های فرکانسی مکانی استخراج شده استفاده شد. این نشان می‌دهد که MFCC می‌تواند به‌عنوان ابزاری برای استخراج ویژگی در تشخیص خرابی یاتاقان استفاده شود. MFCC برای استخراج ویژگی‌های

¹ Convolutional Neural Network

² Feature Extraction

³ Short-time Fourier Transform

⁴ Wavelet Transform

⁵ Acoustic Spectroscopy

⁶ Malta Fairs and Conventions Centre

⁷ White Noise

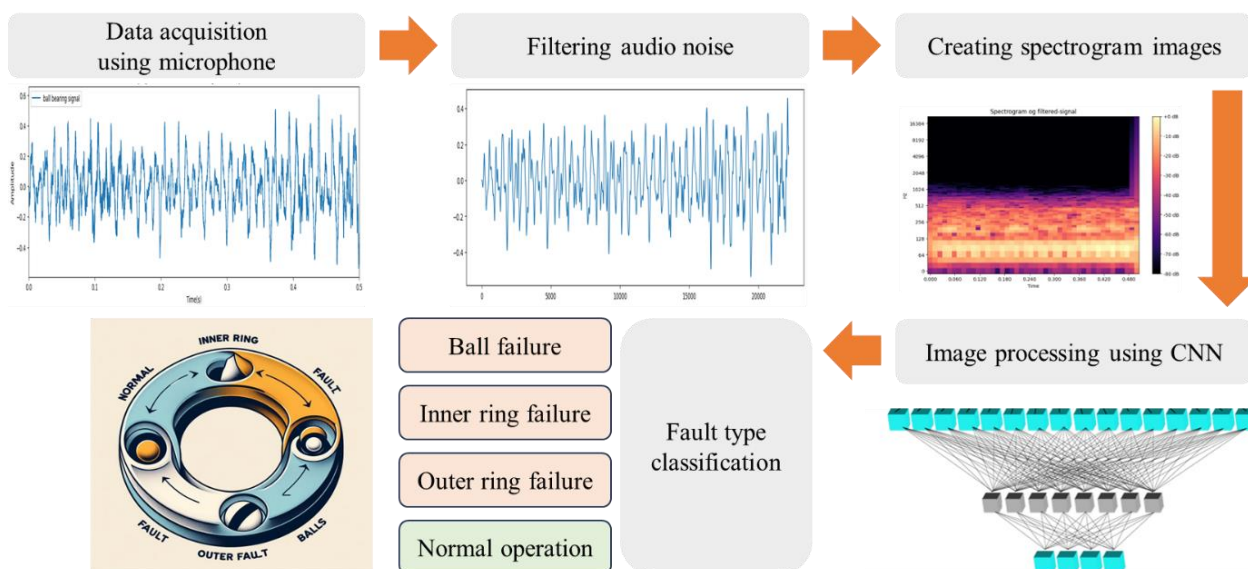
⁸ Cepstral Analysis

سیگنال‌های صوتی مؤثر است. MFCC مورد استفاده در مطالعات موجود نیاز به نرخ نمونه‌برداری بالا دارد و از باند فرکانسی وسیعی استفاده می‌کند، حتی زمانی که از سیگنال‌های ارتعاشی استفاده می‌شود. علاوه بر این، CNN پیچیدگی فرایند طبقه‌بندی خرابی را افزایش می‌دهد. در همین زمینه، یک مطالعه بر روی ویژگی‌های خرابی مشخصه را از باند فرکانسی تشدید سیگنال‌های ارتعاشی استخراج کرده است. باند فرکانسی تشدید یک باند فرکانسی است که نسبتاً بالاتر از فرکانس خرابی یاتاقان است [۲۱]. فام و همکاران [۲۲] به بررسی روش‌های نوین برای تشخیص خرابی یاتاقان‌ها در ماشین‌های دوار می‌پردازد. یاتاقان‌ها از اجزای حیاتی ماشین‌های دوار هستند و خرابی آن‌ها می‌تواند منجر به توقف‌های ناخواسته و هزینه‌های تعمیراتی بالا شود. این پژوهش با استفاده از سیگنال‌های ارتعاشی و STFT به تولید طیف‌نگارهایی می‌پردازد، که به‌عنوان ورودی به CNN داده می‌شوند. مدل پیشنهادی از معماری VGG16 برای استخراج ویژگی‌ها و طبقه‌بندی وضعیت سلامت یاتاقان‌ها استفاده می‌کند. نتایج آزمایش‌ها نشان می‌دهد که این روش دقت و پایداری بالایی در تشخیص خرابی یاتاقان‌ها حتی در محیط‌های پر نویز دارد. این مطالعه نشان داد که استفاده از روش‌های یادگیری عمیق می‌تواند به طور قابل توجهی فرایند تشخیص خرابی را بهبود بخشد و به جایگزینی فرآیندهای دستی توسط مهندسان نگهداری کمک کند.

در این پژوهش، یک سیستم تشخیص خرابی یاتاقان پیشنهاد شده است که شامل سه مرحله است: پیش‌پردازش داده‌ها، استخراج ویژگی و طبقه‌بندی خرابی. روش استخراج ویژگی برای سیستم پیشنهادی، یک طیف‌نگاری با تأکید بر ویژگی‌های فرکانسی صوتی است. سپس، سیستم پیشنهادی طبقه‌بندی خرابی یک مدل CNN بهینه‌سازی شده است که از طیف‌نگاری صوتی به‌عنوان تصویر ورودی استفاده می‌کند. در این مدل از تبدیل کسینوسی گسسته^۱ (DCT) در طبقه‌بندی استفاده نشده است، زیرا اطلاعات ویژگی را کاهش می‌دهد و نیاز به محاسبات اضافی دارد. علاوه بر این، مدل CNN پیشنهادی برای طبقه‌بندی مؤثر بهینه‌سازی و طراحی شده است.

۲- مدل شبکه عصبی طیف‌نگاری

شماتیک کلی سیستم پیشنهادی تشخیص خرابی یاتاقان در شکل ۱ نشان داده شده است. این فرایند شامل داده‌برداری، پیش‌پردازش و پردازش تصویر طیف نگاره جهت طبقه‌بندی نوع خرابی یاتاقان است. داده‌برداری توسط میکروفون ضبط صدا انجام می‌شود و در ادامه پیش‌پردازش، شامل فیلتر نویزهای موجود روی داده‌ها اعمال می‌شود که خروجی مرحله پیش‌پردازش آن یک تصویر طیف نگاره صوتی است. این تصویر در مدل CNN پردازش و نوع خرابی تعیین می‌گردد.



شکل ۱ شماتیک مراحل پردازش سیگنال‌های صوتی جهت طبقه‌بندی نوع خرابی

^۱ Discrete Cosine Transform

فرکانس‌های خرابی یاتاقان‌ها معمولاً به‌صورت دوره‌ای ظاهر می‌شوند. این عیوب عمدتاً در اثر سایش یکی از اجزای یاتاقان به وجود می‌آیند. فرکانس عبور ساچمه از سطح بیرونی^۱ (BPFO) زمانی مشاهده می‌شود که مسیر بیرونی یاتاقان دچار خرابی شده باشد؛ در این حالت، ساچمه هنگام عبور از محل خرابی روی مسیر بیرونی، تولید صدا می‌کند. فرکانس عبور ساچمه از سطح داخلی^۲ (BPI) زمانی ظاهر می‌شود که مسیر داخلی یاتاقان معیوب باشد؛ در این حالت، ساچمه در حال چرخش از روی خرابی عبور کرده و صدا تولید می‌کند. فرکانس چرخش ساچمه^۳ (BSF) نیز زمانی مشاهده می‌شود که خود ساچمه‌ها دارای خرابی باشند؛ در این حالت، ساچمه‌ها ضمن عبور از مسیر یاتاقان، به‌دور محور خود چرخیده و لرزش‌های دوره‌ای ایجاد می‌کنند. تمامی این عیوب باعث تولید لرزش و در نتیجه تولید صدا می‌شوند که منجر به ظهور یک پیک مشخص در حوزه فرکانس می‌گردد. فرکانس خرابی برای انواع مختلف خرابی یاتاقان می‌تواند بسته به مشخصات یاتاقان و سرعت چرخشی ماشین متغیر باشد. روابط ۱ تا ۳ فرکانس‌های خرابی را بر اساس نوع خرابی یاتاقان نشان می‌دهند [۲۳].

$$OUTF = \frac{nf_r}{2} \left\{ 1 + \frac{d}{D} \cos\alpha \right\} \quad (1)$$

$$INNF = \frac{nf_r}{2} \left\{ 1 - \frac{d}{D} \cos\alpha \right\} \quad (2)$$

$$BALLF = \frac{Df_r}{2d} \left\{ 1 - \left(\frac{d}{D} \cos\alpha \right)^2 \right\} \quad (3)$$

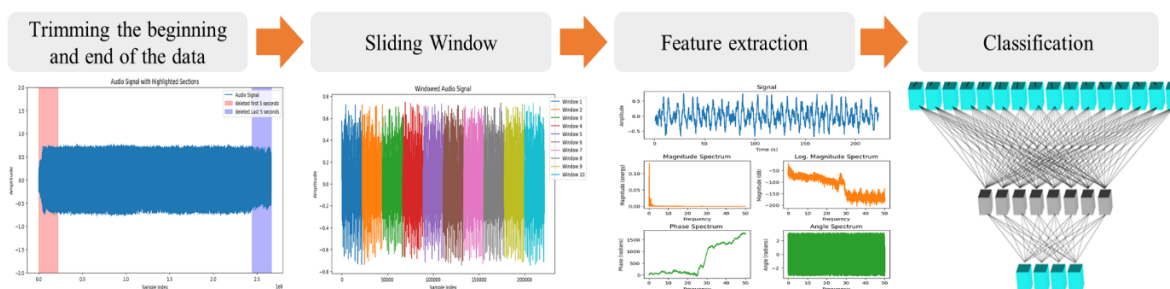
که در آن D ، d ، α ، n و f_r به ترتیب قطر گام، قطر توپ، زاویه تماس بین توپ و قفسه، تعداد عناصر غلتشی و سرعت چرخش یاتاقان (هرتز) هستند. برای مجموعه داده‌های تشخیص خرابی یاتاقان، از مجموعه داده‌های دانشگاه اوتوا استفاده شد [۲۴]. این کتابخانه شامل سیگنال‌های صوتی از یاتاقان‌هایی است که در ماشین‌آلات دوار استفاده می‌شوند. سیگنال‌های صوتی با استفاده از یک میکروفون متصل به ماشین دوار جمع‌آوری شده‌اند.

جدول ۱ پارامترهای ثابت نمونه‌برداری بر اساس پایگاه داده دانشگاه اتاوا

بارگذاری ثابت	سرعت دورانی یاتاقان	زمان داده‌برداری از هر نمونه	فرکانس نمونه‌برداری
400 N	1.750 rpm	10 s	42 kHz

۲-۱- روش آنالیز داده

با توجه به محدودیت‌های معرفی‌شده در حوزه نمونه‌برداری داده‌های صوتی، نیاز است از مرحله نمونه‌برداری تا آموزش شبکه عصبی، عملیات مختلفی صورت بگیرد. روش پیشنهادی شامل ۵ فرایند اصلی است: قاب‌بندی و پنجره‌بندی، فیلترکردن، استخراج ویژگی، ترتیب‌دهی ویژگی‌ها و طبقه‌بندی. این مراحل به‌طور خلاصه در شکل ۲ نشان داده شده است. در فرایند فیلترکردن، برای نمایش سیگنال‌های خرابی یاتاقان‌ها طیف‌نگاری با استفاده از تبدیل فوریه به‌عنوان ویژگی دو کانالی استخراج می‌شوند. در نهایت، با توجه به عملکرد رضایت‌بخش در پردازش سیگنال‌های ۲ بعدی، مانند تشخیص چهره و پلاک خودرو، یک CNN برای برچسب‌گذاری نمونه‌های ورودی ایجاد شده است.



شکل ۲ مراحل پیش‌پردازش داده‌ها و طبقه‌بندی جهت ایجاد ورودی شبکه عصبی

¹ Ball Pass Frequency Outer

² Ball Pass Frequency Inner

³ Ball Spin Frequency

۲-۲- نرمال‌سازی داده‌های صوتی

نرمال‌سازی داده‌های صوتی به معنای تنظیم دامنه سیگنال صوتی به یک محدوده مشخص است. این فرایند معمولاً برای بهبود کیفیت و دقت پردازش سیگنال انجام می‌شود. سیگنال‌های صوتی می‌توانند دارای دامنه‌های مختلفی باشند که ممکن است پردازش آن‌ها را دشوار کند. نرمال‌سازی به استانداردسازی داده‌ها کمک می‌کند تا تمامی سیگنال‌ها در یک دامنه مشخص، معمولاً بین ۰ و ۱ قرار گیرند.

یکی از مهم‌ترین مزایای نرمال‌سازی، کاهش تأثیر نویزهای ناخواسته در داده‌ها است. همچنین در کاربردهای مربوط به یادگیری ماشین، نرمال‌سازی داده‌های ورودی منجر به بهبود دقت مدل‌ها می‌شود. مدل‌های یادگیری ماشین معمولاً با داده‌های نرمال‌شده بهتر عمل می‌کنند، زیرا این داده‌ها دارای مقیاس یکنواختی هستند که باعث افزایش قابلیت تعمیم مدل می‌شود. در این پژوهش، برای نرمال‌سازی داده‌ها از روش نرمال‌سازی مینیمم-ماکزیمم استفاده شده است. این روش داده‌های ورودی را بین صفر و یک مقیاس‌بندی می‌کند تا تمامی مقادیر در یک محدوده ثابت قرار بگیرند. فرمول این روش به صورت زیر بیان می‌شود:

$$Z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (4)$$

که در اینجا (X) بردار داده اولیه و (Z) بردار داده‌های نرمال شده است. این روش باعث یکنواخت‌سازی داده‌ها می‌شود و پردازش آن‌ها را تسهیل می‌کند.

۲-۳- پنجره‌بندی داده‌های صوتی

برای آموزش مدل هوش مصنوعی پیشنهادی، از مجموعه‌ای از فایل‌های صوتی موجود در پایگاه داده دانشگاه اتاوا استفاده شده است. هر فایل شامل ۶۰ ثانیه صدا با فرکانس نمونه‌برداری ۴۴۱۰۰ هرتز است. اما به منظور کاهش خطاهای پردازشی، ۵ ثانیه ابتدایی و انتهایی هر فایل حذف شده است تا داده‌ها یک‌دست شوند.

برای بهبود کیفیت پردازش، یک فیلتر فرکانسی اعمال شده است که فرکانس‌های بالاتر از ۱۲۰۰ هرتز را حذف کرده و فقط فرکانس‌های پایین‌تر را حفظ می‌کند. این کار باعث حذف نویزهای ناخواسته و تمرکز بر اطلاعات کلیدی سیگنال می‌شود. از ۵۰ ثانیه باقیمانده، پنجره‌هایی با طول ۰/۵ ثانیه استخراج شده‌اند. با توجه به فرکانس نمونه‌برداری، هر پنجره شامل ۲۲۵۰ نمونه از داده‌های صوتی است. در مجموع، برای آموزش شبکه عصبی ۴۰۰ پنجره صوتی آماده‌سازی شده است که انواع عیوب رینگ داخلی و خارجی، خرابی تویی‌ها و همچنین حالت نرمال کاری یاتاقان را پوشش می‌دهد.

۲-۴- استخراج ویژگی‌های صوتی

یکی از روش‌های رایج برای تحلیل داده‌های صوتی، طیف‌نگاری صوتی^۱ است. طیف‌نگاری یک نمایش گرافیکی از فرکانس‌های مختلف یک سیگنال صوتی در طول زمان است. این تصاویر به صورت دوبعدی نمایش داده می‌شوند؛ محور افقی نشان‌دهنده زمان و محور عمودی نشان‌دهنده فرکانس‌ها است. شدت رنگ‌ها در این نمایش نشان‌دهنده میزان قدرت یا شدت فرکانس‌های موجود در هر لحظه است.

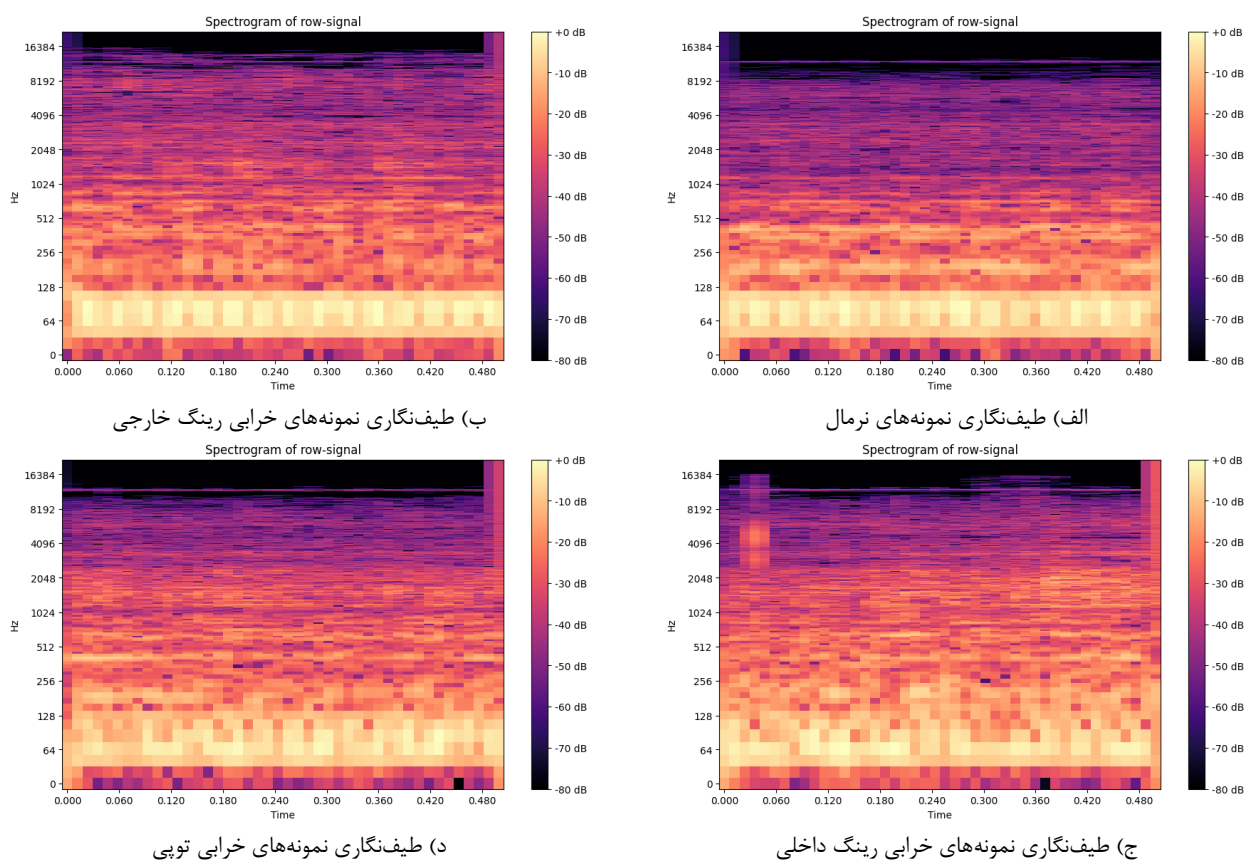
برای ساخت طیف‌نگاری، ابتدا سیگنال صوتی به بخش‌های کوچکی تقسیم می‌شود. سپس برای هر بخش، تبدیل فوریه سریع^۲ (FFT) انجام شده تا فرکانس‌های موجود در آن بخش مشخص شوند. شکل ۳ طیف‌نگارهای مربوط به حالت‌های مختلف خرابی یاتاقان را نشان می‌دهد. تبدیل فوریه یک ابزار ریاضی قدرتمند است که سیگنال را از حوزه زمان به حوزه فرکانس تبدیل می‌کند. فرمول این تبدیل به صورت رابطه شماره ۵ بیان می‌شود:

$$x(t) = \sum_{n=0}^{\frac{N}{2}-1} x_{2n} e^{-j2\pi\omega n / (\frac{N}{2})} + e^{-j2\pi\omega n / (\frac{N}{2})} \sum_{n=0}^{\frac{N}{2}-1} x_{2n+1} e^{-j2\pi\omega n / (\frac{N}{2})} \quad (5)$$

¹ Spectrogram

² Fast Fourier Transform

در این رابطه، $x(t)$ سیگنال در حوزه زمان و $X(f)$ سیگنال در حوزه فرکانس است. با استفاده از این تبدیل، می‌توان ویژگی‌های مهم سیگنال‌های صوتی را استخراج کرده و برای تحلیل و پردازش در مدل‌های یادگیری ماشین استفاده کرد.



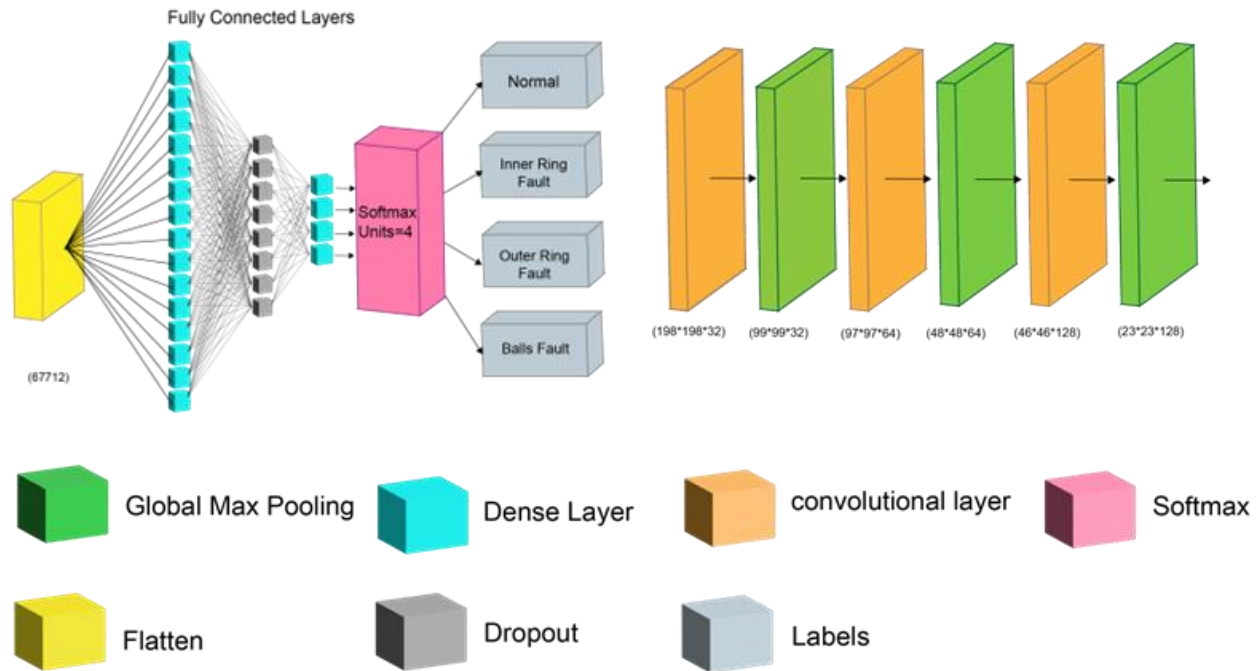
شکل ۳ طیف‌نگاری صوت در حالت کاری مختلف

۲-۵- طبقه بند کننده

مدل پیشنهادی یک CNN بهینه‌سازی شده برای دستیابی به طبقه‌بندی دقیق تصاویر طراحی شده است. جدول ۲ جزئیات لایه‌های مدل CNN را نشان می‌دهد، شکل ۴ نیز نمایی بصری از پیکربندی بهینه‌شده مدل CNN را نشان می‌دهد.

جدول ۲ پارامترهای ثابت نمونه‌برداری

شماره لایه	نام لایه	اندازه داده در لایه	اندازه فیلتر لایه	تابع عملگر
۱	Conv2D	(32, 708, 298)	(3,3,32)	Relu
۲	MaxPooling2D	(32, 354, 149)	(2,2)	-
۳	Conv2D	(64, 352, 147)	(3,3,64)	Relu
۴	MaxPooling2D	(64, 176, 73)	(2,2)	-
۵	Conv2D	(128, 174, 71)	(3,3,128)	Relu
۶	MaxPooling2D	(128, 87, 35)	(2,2)	-
۷	Flatten	(389760)	-	-
۸	Dense	(128)	-	Relu
۹	Dropout	(128)	(0.5)	-
۱۰	Dense	(128)	-	Relu
۱۱	Dense	(4)	-	softmax



شکل ۴ شماتیک معماری شبکه CNN پیشنهادی

در ادامه لایه‌های شبکه تشریح شده‌اند.

- لایه اول (Conv2D) یک‌لایه دوبعدی است که ۳۲ فیلتر با اندازه 3×3 اعمال می‌کند. ورودی این لایه، یک تصویر طیف‌نگاره‌ای با اندازه 30×30 و ۳ کانال رنگی است. در این لایه و لایه‌های ماقبل آخر تابع فعال‌سازی ReLU برای یادگیری ویژگی‌های غیرخطی استفاده می‌گردد.
- لایه دوم (MaxPooling2D) عملیات نمونه‌برداری حداکثر را با اندازه 2×2 انجام می‌دهد که اندازه ویژگی‌های استخراج‌شده را کاهش می‌دهد و محاسبات را بهینه می‌کند.
- لایه سوم (Conv2D) ۶۴ فیلتر با اندازه 3×3 اعمال می‌کند و از تابع فعال‌سازی ReLU استفاده می‌کند.
- لایه چهارم (MaxPooling2D) نیز عملیات نمونه‌برداری حداکثر را با اندازه 2×2 انجام می‌دهد.
- لایه پنجم (Conv2D) ۱۲۸ فیلتر با اندازه 3×3 اعمال می‌کند و از تابع فعال‌سازی ReLU استفاده می‌کند.
- لایه آخر (MaxPooling2D) عملیات نمونه‌برداری حداکثر را با اندازه 2×2 انجام می‌دهد.
- لایه Flatten تمام ویژگی‌های دوبعدی را به یک بردار یک‌بعدی تبدیل می‌کند تا بتوان آن‌ها را به لایه‌های کاملاً متصل داد. لایه‌های کاملاً متصل شامل لایه‌های Dense و Dropout است که این لایه به طور تصادفی ۵۰٪ از نورون‌ها را در هر مرحله آموزشی غیرفعال می‌کند تا از بیش‌برازش جلوگیری کند و تعمیم‌پذیری مدل را بهبود بخشد. در لایه Dense نهایی خروجی ۴ نورون دارد و از تابع فعال‌سازی SoftMax استفاده می‌کند که احتمال تعلق هر ورودی به یکی از ۴ کلاس را محاسبه می‌کند. این معماری کمک می‌کند تا تصاویر ورودی به طور مؤثری طبقه‌بندی شوند و با استفاده از لایه‌های پیچشی و نمونه‌برداری حداکثر، ویژگی‌های مهم استخراج و اندازه داده‌ها کاهش می‌یابد. لایه‌های کاملاً متصل و SoftMax نیز برای طبقه‌بندی نهایی استفاده می‌شوند.

۳- نتایج و بحث

سیستم تشخیص پیشنهادی با استفاده از مجموعه داده دانشگاه اتاوا ارزیابی شده است. این مجموعه داده شامل یاتاقان‌هایی است که به انتهای درایو و فن متصل و با استفاده از میکروفون اندازه‌گیری شده‌اند. کلاس‌های تشخیص خرابی یاتاقان به سیگنال‌های صوتی در حالت عادی به‌عنوان حالت کارکرد عادی و سیگنال‌های صوتی خرابی رینگ خارجی، خرابی رینگ داخلی و خرابی توپ به‌عنوان کلاس‌های خرابی اختصاص داده شده‌اند. سیگنال‌های صوتی با نرخ نمونه‌برداری ۴۴۱۰۰ هرتز ارائه شده و به‌صورت فایل wav

ذخیره‌شده است. روی داده‌های خام، فیلتر حذف نویز اعمال شد تا تأثیرات نویزهای کوچک از بین برود. سیگنال صوتی از ثانیه ۵ شروع و تا ثانیه ۵۵ در نظر گرفته شده و نقطه شروع تکرار نشده است. شناسایی خرابی‌های یاتاقان به شرکت‌های تولید و تعمیر یاتاقان کمک می‌کند تا خرابی‌های آینده را تحلیل و پیش‌بینی کنند. کارخانه‌های صنعتی که از یاتاقان‌ها استفاده می‌کنند، باید بتوانند به‌سرعت خرابی این قطعات را تشخیص دهند تا از خرابی تجهیزات اساسی جلوگیری کنند و یاتاقان‌ها باید بدون توجه به نوع خرابی تعویض شوند. در مدل CNN پیشنهادی با لایه‌های معرفی‌شده داده‌های آزمون با ۳۰ تکرار آموزش داده شده است. که حجم مدل نهایی آن ۳۳,۴۲ مگابایت گزارش شده و دقت نهایی حاصل شده ۰/۹۹۸۷ است. در جدول ۳ جزئیات مربوط به آموزش مدل آورده شده است.

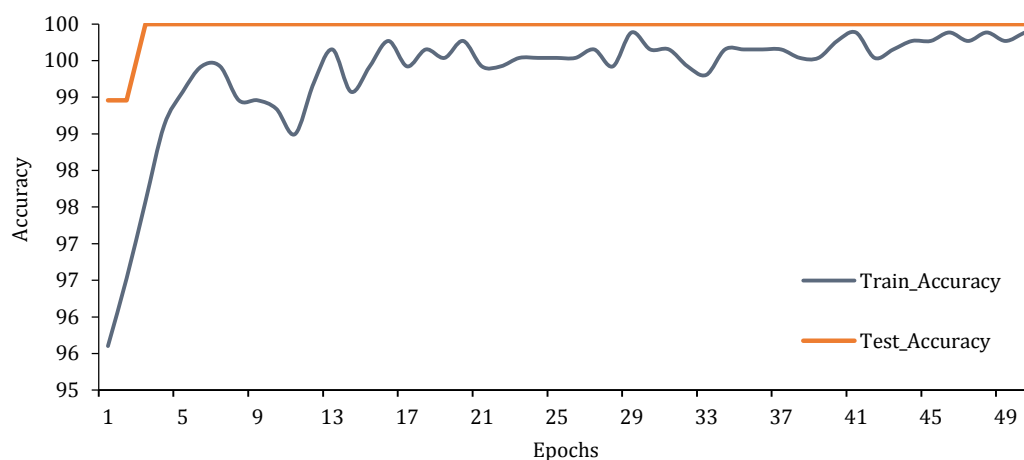
جدول ۳ پارامترهای مدل CNN پیشنهادی

اندازه نمونه	نرخ یادگیری	صحت	خطا
25	0.001	0.9987	0.0001

برای ارزیابی دقت مدل پیشنهادی برای مجموعه داده Ottawa، دقت به‌عنوان تابعی از ضریب شیب و دوره‌های آزمون مدل بررسی شده است. جدول ۴ دقت داده‌های آزمون و دقت مدل روی داده‌های آزمون را نشان می‌دهد و از طریق این تحلیل، دقت^۱ و تضرر^۲ طبقه‌بندی خرابی یاتاقان به‌دست آمده است. نتایج نشان داد که بالاترین دقت و تضرر داده‌های آزمون به ترتیب ۰/۹۹۸۸ و ۰/۰۱۶۳ بوده است. شکل ۶ فرایند آموزش را نشان می‌دهد که برای به‌دست آوردن دقت طبقه‌بندی و تضرر داده‌های آموزش و آزمون مدل پیشنهادی استفاده شده است، درحالی‌که جدول ۴ دقت و تضرر در آموزش و آزمون را نشان می‌دهد. از آنجاکه داده‌های آموزش و آزمون از یک آموزش به آموزش دیگر تغییر می‌کنند، دقت و تضرر تکنیک پیشنهادی نشان‌دهنده پایداری است.

جدول ۴ مدل CNN پیشنهادی

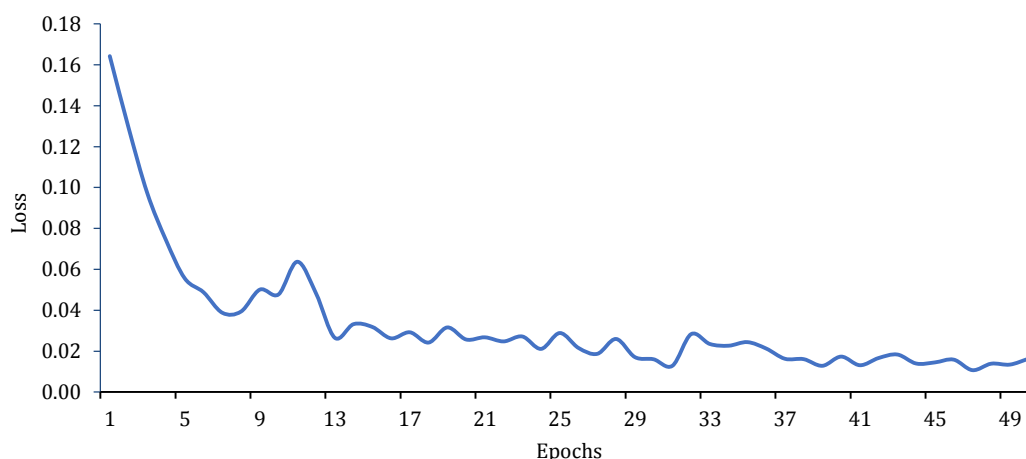
معیار تضرر روی داده‌های آزمون	دقت داده‌های آزمون	شماره تکرار آزمون
0.0556	99.07	5
0.0478	98.84	10
0.0318	99.42	15
0.0257	99.77	20
0.0289	99.54	25
0.0160	99.65	30
0.0244	99.65	35
0.0173	99.77	40
0.0145	99.77	45
0.0163	99.88	50



شکل ۵ فرایند آموزش مدل CNN

¹ Accuracy² Loss

در این مطالعه، مشاهده شد که در دوره ۱۶ دقت مدل به بالاترین حد ممکن رسید و در این نقطه ثابت ماند و همچنین در شکل ۶ مشاهده می‌شود که تضرر در داده‌های آزمون با سرعت بیشتری کاهش می‌شود.



شکل ۶ کاهش تضرر در دوره مدل CNN

۴- نتیجه‌گیری

در این پژوهش، یک سیستم تشخیص خرابی یاتاقان پیشنهاد گردید که فرایند طیف‌نگاری صوتی و یک شبکه CNN را با استفاده از اندازه‌گیری تماس سیگنال‌های ارتعاشی ترکیب می‌کند. این الگوریتم، ابتدا سیگنال صوتی را به یک تصویر طیف‌نگاری صوتی با ابعاد $3 \times 30 \times 3$ تبدیل نمود. تصویر تولیدشده به‌عنوان ورودی به CNN برای طبقه‌بندی خرابی در بلبرینگ استفاده گردید. در ادامه استخراج ویژگی طیف‌نگاری صوتی و پیچیدگی CNN ارزیابی شد. نتایج نشان از دقت طیف‌نگار $99/88\%$ داشت که بالاترین دقت در میان مجموعه تصاویر تولیدشده است. مدل CNN دارای $622/77$ کیلوبایت پارامتر و تعداد عملیات ممیز شناور در ثانیه $1/53 \times 10^6$ است. ارزیابی نشان داد که مشکلات نرخ نمونه‌برداری بالا، استفاده زیاد از تحلیل فرکانس پهن باند موجود و پیچیدگی بالای مدل‌های شبکه عصبی عمیق حل شده است. به‌ویژه، فرایند طیف‌نگاری صوتی می‌تواند باند فرکانس را با تغییر ضریب شیب کنترل کند. بنابراین، فرایند طیف‌نگاری صوتی حتی زمانی که فرکانس چرخش یاتاقان تغییر می‌کند، قابل اعمال است.

References

- [1] Grebenik J, Zhang Y, Bingham C, Srivastava S. Roller element bearing acoustic fault detection using smartphone and consumer microphones comparing with vibration techniques. In 2016 17th International Conference on Mechatronics-Mechatronika (ME) 2016 Dec 7 (pp. 1-7). IEEE.
- [2] Rasti A, Jafari D, Rafierad I. Design and performance analysis of integrated roller bearing and roller screw mechanism for precision linear motion. Iranian Journal of Manufacturing Engineering. 2025 Apr 21;12(2):59-69. [in Persian] doi: 10.22034/ijme.2025.495859.2038
- [3] Ramezani NM, Rasti A, Sadeghi MH, Pour BJ, Hajideh MR. Experimental study of tool wear and surface roughness on high speed helical milling in D2 steel. Modares Mech Eng. 2016;15(20):198-202. [in Persian]
- [4] Hassanpour H, Rasti A, Sadeghi MH, Khosrowshahi JH. Investigation of roughness, topography, microhardness, white layer and surface chemical composition in high speed milling of Ti-6Al-4V using minimum quantity lubrication. Machining Science and Technology. 2020 Sep 2;24(5):719-38. doi: 10.1080/10910344.2020.1752237
- [5] Zeinolabedin-Beygi A, Naeini HM, Talebi-Ghadikolaee H, Rabiee AH, Hajiahmadi S. Predictive modeling of spring-back in pre-punched sheet roll forming using machine learning. The Journal of Strain Analysis for Engineering Design. 2024 Oct;59(7):463-74. doi: 10.1177/03093247241263685
- [6] Hajiahmadi S, Moslemi Naeini H, Talebi-Ghadikolaee H, Safdarian R, Zeinolabedin-Beygi A. Insights into spring-back prediction: a comparative analysis of constitutive models for perforated U-shaped roll-formed steel profiles. The International Journal of Advanced Manufacturing Technology. 2024 Sep;134(3):1915-33. doi: 10.1007/s00170-024-14211-5

- [7] Zhang S, Zhang S, Wang B, Habetler TG. Deep learning algorithms for bearing fault diagnostics—A comprehensive review. IEEE access. 2020 Feb 10;8:29857-81. doi: [10.1109/ACCESS.2020.2972859](https://doi.org/10.1109/ACCESS.2020.2972859)
- [8] Abdenebi R, Fethi D, Abdelkrim N, Badis D. Gearbox fault diagnosis using the short-time cepstral features. In 2022 2nd International Conference on Advanced Electrical Engineering (ICAEE) 2022 Oct 29 (pp. 1-8). IEEE. doi: [10.1109/ICAEE53772.2022.9962139](https://doi.org/10.1109/ICAEE53772.2022.9962139)
- [9] Kumar A, Vashishtha G, Gandhi CP, Zhou Y, Glowacz A, Xiang J. Novel convolutional neural network (NCNN) for the diagnosis of bearing defects in rotary machinery. IEEE Transactions on Instrumentation and Measurement. 2021 Feb 1;70:1-0. doi: [10.1109/TIM.2021.3055802](https://doi.org/10.1109/TIM.2021.3055802)
- [10] Azeez AA, Alkhedher M, Gadala MS. Thermal imaging fault detection for rolling element bearings. In 2020 Advances in science and engineering technology international conferences (ASET) 2020 Feb 4 (pp. 1-5). IEEE. doi: [10.1109/ASET48392.2020.9118361](https://doi.org/10.1109/ASET48392.2020.9118361)
- [11] Liu H, Li L, Ma J. Rolling bearing fault diagnosis based on STFT-deep learning and sound signals. Shock and Vibration. 2016;2016(1):6127479. doi: [10.1155/2016/6127479](https://doi.org/10.1155/2016/6127479)
- [12] Sun D, Fan Y, Wang G. Gradient-oriented prioritization in meta-learning for enhanced few-shot fault diagnosis in industrial systems. Applied Sciences. 2023 Dec 25;14(1):181. doi: [10.3390/app14010181](https://doi.org/10.3390/app14010181)
- [13] Shan S, Liu J, Wu S, Shao Y, Li H. A motor bearing fault voiceprint recognition method based on Mel-CNN model. Measurement. 2023 Feb 15;207:112408. doi: [10.1016/j.measurement.2022.112408](https://doi.org/10.1016/j.measurement.2022.112408)
- [14] Elhami S, Razfar MR, Farahnakian M, Rasti A. Application of GONNS to predict constrained optimum surface roughness in face milling of high-silicon austenitic stainless steel. The International Journal of Advanced Manufacturing Technology. 2013 May;66(5):975-86. doi: [10.1007/s00170-012-4382-y](https://doi.org/10.1007/s00170-012-4382-y)
- [15] Li D, Li B, Wang C, Cheng P, Jiao B. A fault diagnosis equipment of motor bearing based on sound signal and CNN. In Journal of Physics: Conference Series 2021 Sep 1 (Vol. 2010, No. 1, p. 012159). IOP Publishing. doi: [10.1088/1742-6596/2010/1/012159](https://doi.org/10.1088/1742-6596/2010/1/012159)
- [16] Wang H, Liu Z, Peng D, Cheng Z. Attention-guided joint learning CNN with noise robustness for bearing fault diagnosis and vibration signal denoising. ISA transactions. 2022 Sep 1;128:470-84. doi: [10.1016/j.isatra.2021.11.028](https://doi.org/10.1016/j.isatra.2021.11.028)
- [17] Wang X, Mao D, Li X. Bearing fault diagnosis based on vibro-acoustic data fusion and 1D-CNN network. Measurement. 2021 Mar 1;173:108518. doi: [10.1016/j.measurement.2020.108518](https://doi.org/10.1016/j.measurement.2020.108518)
- [18] Chen J, Jiang J, Guo X, Tan L. An Efficient CNN with Tunable Input-Size for Bearing Fault Diagnosis. Int. J. Comput. Intell. Syst.. 2021 Jan 1;14(1):625-34. doi: [10.2991/ijcis.d.210113.001](https://doi.org/10.2991/ijcis.d.210113.001)
- [19] Xia M, Li T, Xu L, Liu L, De Silva CW. Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks. IEEE/ASME transactions on mechatronics. 2017 Jul 17;23(1):101-10. doi: [10.1109/TMECH.2017.2728371](https://doi.org/10.1109/TMECH.2017.2728371)
- [20] Akpudo UE, Hur JW. A cost-efficient MFCC-based fault detection and isolation technology for electromagnetic pumps. Electronics. 2021 Feb 10;10(4):439. doi: [10.3390/electronics10040439](https://doi.org/10.3390/electronics10040439)
- [21] Zhao Y, Zhang N, Zhang Z, Xu X. Bearing fault diagnosis based on mel frequency cepstrum coefficient and deformable space-frequency attention network. IEEe Access. 2023 Apr 3;11:34407-20. doi: [10.1109/ACCESS.2023.3264276](https://doi.org/10.1109/ACCESS.2023.3264276)
- [22] Pham MT, Kim JM, Kim CH. Deep learning-based bearing fault diagnosis method for embedded systems. Sensors. 2020 Dec 2;20(23):6886. doi: [10.3390/s20236886](https://doi.org/10.3390/s20236886)
- [23] Mohamad TH. Nonlinear Dynamics-Based Machine Learning Algorithms for Condition-Based Maintenance. Villanova University; 2021.
- [24] Sehri M, Dumond P, Bouchard M. University of Ottawa constant load and speed rolling-element bearing vibration and acoustic fault signature datasets. Data in Brief. 2023 Aug 1;49:109327. doi: [10.1016/j.dib.2023.109327](https://doi.org/10.1016/j.dib.2023.109327)