

Geliş Tarihi : 27/05/2026
Kabul Tarihi :12/06/2026

Received Date :27/05/2026
Accepted Date :12/06/2026

RULMAN ARIZA TEŞHİSİ İÇİN CNN-LSTM VE CNN-GRU HİBRİT DERİN ÖĞRENME MODELLERİNİN KARŞILAŞTIRMALI BİR ÇALIŞMASI

A COMPARATIVE STUDY OF CNN-LSTM AND CNN-GRU HYBRID DEEP LEARNING MODELS FOR BEARING FAULT DIAGNOSIS

Eyyüp AKCAN^{1*} (ORCID: 0000-0002-4133-4344)
Hasan EROĞLU² (ORCID: 0000-0001-5644-4178)

¹Siirt Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Motorlu Araçlar ve Ulaştırma Teknolojileri Bölümü, Siirt, Türkiye
²Siirt Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Siirt, Türkiye

*Sorumlu Yazar / Corresponding Author: Eyyüp AKCAN

Öz

Rulman arıza teşhisi, endüstriyel döner makinelerin güvenli ve verimli çalışmasının sağlanması açısından kritik bir öneme sahiptir. Bu çalışmada, rulman arızalarının sınıflandırılması amacıyla iki farklı hibrit derin öğrenme modeli olan Evrişimsel Sinir Ağı-Uzun Kısa Süreli Bellek (CNN-LSTM) ve Evrişimsel Sinir Ağı-Geçitli Tekrarlayan Birim (CNN-GRU) modelleri önerilmiş ve karşılaştırmalı olarak analiz edilmiştir. Çalışmada, Case Western Reserve University (CWRU) rulman veri setinden elde edilen titreşim sinyalleri kullanılmıştır. CNN yapısı, ham sinyallerden otomatik özellik çıkarımı için kullanılırken, LSTM ve GRU katmanları ise verilerin zamansal bağımlılıklarını modellemek amacıyla kullanılmıştır. Önerilen modellerin performansı doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1-skoru metrikleri kullanılarak değerlendirilmiştir. Elde edilen sonuçlara göre, CNN-LSTM modeli %96.32 test doğruluğu elde ederken, CNN-GRU modeli %97.79 doğruluk ile daha yüksek bir performans sergilemiştir. Karışıklık matrisleri ve sınıf bazlı analizler, CNN-GRU modelinin özellikle zor ayrıştırılan arıza sınıflarında daha kararlı ve başarılı sonuçlar ürettiğini göstermektedir. Sonuç olarak, hibrit derin öğrenme modellerinin rulman arıza teşhisi problemlerinde oldukça etkili olduğu ve CNN-GRU modelinin daha basit yapısı ve güçlü zamansal öğrenme yeteneği sayesinde daha üstün performans sağladığı belirlenmiştir. Bu çalışma, endüstriyel durum izleme sistemlerinde güvenilir ve akıllı arıza teşhis yöntemlerinin geliştirilmesine katkı sunmaktadır.

Anahtar Kelimeler: Rulman Arıza Teşhisi, Derin Öğrenme, CNN-LSTM, CNN-GRU, Titreşim Analizi

Abstract

Bearing fault diagnosis is of critical importance for ensuring the safe and efficient operation of industrial rotating machinery. In this study, two different hybrid deep learning models—Convolutional Neural Network–Long Short-Term Memory (CNN-LSTM) and Convolutional Neural Network–Gated Recurrent Unit (CNN-GRU)—are proposed and comparatively analyzed for the classification of bearing faults. Vibration signals obtained from the Case Western Reserve University (CWRU) bearing dataset are utilized in the study. While the CNN structure is employed for automatic feature extraction from raw signals, the LSTM and GRU layers are used to model the temporal dependencies of the data. The performance of the proposed models is evaluated using accuracy, precision, recall, and F1-score metrics. According to the obtained results, the CNN-LSTM model achieves a test accuracy of 96.32%, whereas the CNN-GRU model demonstrates superior performance with an accuracy of 97.79%. Confusion matrices and class-wise analyses indicate that the CNN-GRU model produces more stable and accurate results, particularly for fault classes that are difficult to distinguish. In conclusion, hybrid deep learning models are highly effective for bearing fault diagnosis problems, and the CNN-GRU model provides better performance due to its simpler architecture and strong capability in modeling temporal dependencies. This study contributes to the development of reliable and intelligent fault diagnosis methods for industrial condition monitoring systems.

Keywords: Bearing Fault Diagnosis, Deep Learning, CNN-LSTM, CNN-GRU, Vibration Analysis

1.GİRİŞ

Döner makinelerde yaygın olarak kullanılan yuvarlanma rulmanları, makine sistemlerinin en kritik bileşenlerinden biri olup, sistemin genel güvenilirliği ve operasyonel güvenliği üzerinde doğrudan belirleyici bir etkiye sahiptir. Bu bileşenlerde meydana gelen herhangi bir arıza, yalnızca ekipmanın performansını düşürmekle kalmayıp aynı zamanda ciddi ekonomik kayıplara, üretim kesintilerine ve hatta insan güvenliğini tehdit eden kazalara yol açabilmektedir. Bu nedenle, döner makinelerde rulmanların durumunun doğru ve zamanında izlenmesi ile arıza teşhisinin güvenilir bir şekilde gerçekleştirilmesi, modern endüstriyel sistemler için hayati bir gereklilik haline gelmiştir. Özellikle son yıllarda, artan üretim kapasitesi ve otomasyon sistemlerinin yaygınlaşması, arıza teşhisi süreçlerinin daha hassas, hızlı ve otomatik hale getirilmesini zorunlu kılmıştır (Huo ve ark., 2023).

Modern endüstride döner makinelerin kullanım alanı giderek genişlemekte olup, bu makinelerin güvenli ve kesintisiz çalışması, bakım maliyetlerinin azaltılması ve sistem verimliliğinin artırılması açısından büyük önem taşımaktadır. Yuvarlanma rulmanları, bu makinelerin düzgün çalışmasını sağlayan temel mekanik elemanlar arasında yer almakta ve rulmanların çalışma durumu, makinenin genel performansını doğrudan etkilemektedir. Bu nedenle, rulman arıza teşhisi konusu, hem akademik hem de endüstriyel araştırmalarda önemli bir çalışma alanı olarak öne çıkmaktadır. Ancak geleneksel arıza teşhis yöntemleri, genellikle manuel özellik çıkarımına ve uzman bilgisine dayalı olması nedeniyle, artan veri hacmi ve karmaşık çalışma koşulları karşısında yetersiz kalmaktadır (Wu ve ark., 2023).

Rulmanlar, ağır yükler, yüksek sıcaklıklar, değişken hızlar ve uzun süreli çalışma gibi zorlu koşullar altında faaliyet gösterdiğinden, titreşim, aşınma, darbe ve erozyon gibi çeşitli fiziksel etkiler nedeniyle arızaya oldukça yatkındır. Bu durum, rulmanları mekanik sistemlerin en zayıf halkalarından biri haline getirmektedir. Özellikle endüstriyel ortamlarda elde edilen titreşim sinyalleri, yoğun gürültü etkisi altında olduğundan, arıza belirtilerinin tespit edilmesi oldukça zorlaşmaktadır. Bu bağlamda, sinyal işleme teknikleri kullanılarak zaman, frekans ve zaman-frekans alanlarında özellik çıkarımı yapılması, arıza teşhisinde önemli bir rol oynamaktadır. Short Time Fourier Transform (STFT), Dalga Dönüşümü (WT), Ampirik Mod Ayırıştırma (EMD) ve Variasyonel Mod Ayırıştırma (VMD) gibi yöntemler, gürültülü sinyallerden anlamlı özelliklerin elde edilmesinde yaygın olarak kullanılmaktadır (Li ve ark., 2023). Bununla birlikte, geleneksel sinyal işleme yöntemleri genellikle uzman bilgisine bağımlıdır ve farklı veri setleri için özel ayarlamalar gerektirir. Bu durum, yöntemlerin genelleme yeteneğini sınırlamakta ve farklı çalışma koşullarında performansın düşmesine neden olmaktadır. Ayrıca, endüstride elde edilen büyük hacimli verilerin manuel olarak analiz edilmesi oldukça zor ve maliyetlidir. Bu nedenle, veri odaklı yaklaşımların geliştirilmesi kaçınılmaz hale gelmiştir. Büyük veri çağının başlamasıyla birlikte, makine öğrenimi ve derin öğrenme tabanlı yöntemler, rulman arıza teşhisi alanında önemli bir alternatif olarak ortaya çıkmıştır (Zhang ve ark., 2023).

Makine öğrenimi tabanlı yöntemler, titreşim sinyallerinden elde edilen özellikleri kullanarak arıza sınıflandırması yapabilmektedir. Destek Vektör Makineleri (SVM), k-En Yakın Komşu (KNN) ve Temel Bileşen Analizi (PCA) gibi geleneksel algoritmalar, bu alanda başarılı sonuçlar elde etmiştir. Ancak bu yöntemlerin performansı, büyük ölçüde elde çıkarılan özelliklerin kalitesine bağlıdır ve karmaşık sistemlerde yeterli doğruluğa ulaşmakta zorlanmaktadır. Bu sınırlamaların aşılması amacıyla, derin öğrenme yöntemleri, otomatik özellik çıkarma yetenekleri sayesinde dikkat çekmektedir. Derin sinir ağları, ham veriden doğrudan anlamlı özellikler öğrenebilmekte ve arıza teşhis sürecini uçtan uca otomatik hale getirebilmektedir (Wang ve ark., 2023).

Derin öğrenme yöntemleri, özellikle büyük veri setlerinin analizinde yüksek doğruluk ve güçlü genelleme yeteneği sunmaktadır. Evrişimli Sinir Ağları (CNN), zaman serisi verilerinden yerel özellikleri çıkarma konusunda oldukça etkilidir. Tekrarlayan Sinir Ağları (RNN) ve bunların gelişmiş versiyonları olan Uzun Kısa Süreli Bellek (LSTM) ve Kapılı Tekrarlayan Birimler (GRU), zaman bağımlılıklarını modelleme konusunda üstün performans göstermektedir. Bu nedenle, CNN ve RNN tabanlı hibrit modeller, hem uzamsal hem de zamansal bilgiyi birlikte kullanarak daha etkili arıza teşhisi sağlamaktadır (Hou ve ark., 2023).

Öte yandan, gerçek endüstriyel uygulamalarda veri setleri genellikle dengesizdir ve bazı arıza türlerine ait veriler sınırlı sayıda bulunmaktadır. Bu durum, modelin genelleme performansını olumsuz etkileyebilmektedir. Ayrıca, arıza verilerinin etiketlenmesi hem zor hem de maliyetli bir süreçtir. Bu nedenle, transfer öğrenme ve veri artırma teknikleri gibi yeni yaklaşımlar geliştirilmiş ve bu alanda önemli ilerlemeler kaydedilmiştir (Gu ve ark., 2023).

Son yıllarda, transfer öğrenme ve derin transfer öğrenme yöntemleri, farklı çalışma koşulları altında elde edilen veriler arasında bilgi transferi sağlayarak model performansını artırmayı hedeflemektedir. Bu yöntemler, farklı yük, hız ve çalışma koşullarında elde edilen veriler arasında genelleme yapabilme yeteneği sunmaktadır. Ayrıca, derin öğrenme modellerinin karmaşık yapısı sayesinde, doğrusal olmayan ve durağan olmayan sinyallerden daha anlamlı özellikler elde edilebilmektedir (Tang ve ark., 2023).

Rulman arıza teşhisi alanında kullanılan veri türleri arasında titreşim sinyalleri, en yaygın ve en güvenilir bilgi kaynağı olarak kabul edilmektedir. Titreşim sinyalleri, arıza durumlarını doğrudan yansıtan karakteristik özellikler içermekte ve kolayca elde edilebilmektedir. Bununla birlikte, bu sinyaller genellikle gürültü, doğrusal olmayanlık ve durağan olmama gibi özellikler taşımaktadır. Bu durum, etkili özellik çıkarımı ve doğru sınıflandırma için önemli bir zorluk oluşturmaktadır (Zhao ve ark., 2023).

Bu çalışmada, rulman arıza teşhisi problemini ele almak amacıyla, evrişimli sinir ağları ile tekrarlayan sinir ağlarının avantajlarını birleştiren hibrit derin öğrenme modelleri incelenmiştir. Özellikle CNN-LSTM ve CNN-GRU mimarileri kullanılarak, titreşim sinyallerinden hem yerel hem de zamansal özelliklerin etkin bir şekilde çıkarılması hedeflenmiştir. Bu kapsamda, önerilen modellerin performansı karşılaştırmalı olarak değerlendirilmiş ve hangi yapının daha yüksek doğruluk ve kararlılık sağladığı analiz edilmiştir.

2. LİTERATÜR ÇALIŞMASI

Yi ve ark. (2025), döner makinelerde arıza teşhisi için derin öğrenme tabanlı hafif ve yüksek performanslı modellerin geliştirilmesine yönelik kapsamlı bir çalışma gerçekleştirmiştir. Bu çalışmada önerilen “VibrMamba” modeli, Mamba mimarisine dayalı olup, zaman serisi titreşim sinyallerini doğrudan işleyebilen iki yönlü durum uzayı modelleri (State Space Models - SSMs) ile Kolmogorov-Arnold ağlarını (KAN) bir araya getiren hibrit bir yapı sunmaktadır. Bu yaklaşım, hem zamansal bağımlılıkların etkin bir şekilde modellenmesini hem de modelin hesaplama verimliliğinin artırılmasını hedeflemektedir. Modelin performansı, CWRU, XJTU, HIT ve PHM-Beijing veri kümeleri üzerinde gerçekleştirilen toplam altı farklı sınıflandırma görevi ile değerlendirilmiştir. Elde edilen bulgular, VibrMamba modelinin hem klasik 1D Mamba yapısına hem de diğer güncel derin öğrenme yöntemlerine kıyasla daha üstün performans sergilediğini ortaya koymuştur. Özellikle, CWRU veri kümesinde %99.77, XJTU veri kümesinde %99.95 ve HIT veri kümesinde iki kanallı giriş kullanıldığında %99.49 doğruluk oranlarına ulaşılması, modelin yüksek doğruluk seviyesine sahip olduğunu göstermektedir. Bununla birlikte, önerilen modelin yalnızca doğruluk açısından değil, aynı zamanda hesaplama maliyeti ve model hafifliği açısından da avantaj sunduğu ifade edilmiştir. Bu yönüyle VibrMamba, özellikle gerçek zamanlı ve kaynak kısıtlı endüstriyel uygulamalar için dikkat çekici ve uygulanabilir bir çözüm olarak değerlendirilmektedir.

Bununla birlikte, Noussis ve ark. (2024), rulman arıza teşhisi amacıyla minimal ön işleme gerektiren ve ham titreşim verilerini doğrudan işleyebilen sade bir çift yönlü uzun-kısa süreli bellek (Bi-LSTM) modeli önermiştir. Çalışmada, Case Western Reserve University (CWRU) rulman veri seti kullanılarak 27 farklı sağlık durumu sınıfı üzerinde kapsamlı deneyler gerçekleştirilmiştir. Önerilen model, zaman serisi verilerini belirli uzunluklarda örnekleyerek tek bir Bi-LSTM katmanı aracılığıyla işlemekte ve ardından tam bağlantılı katman ile sınıflandırma gerçekleştirmektedir. Modelin hiperparametreleri ayrıntılı bir şekilde optimize edilmiş ve aşırı öğrenme problemini azaltmak amacıyla erken durdurma (early stopping) yöntemi uygulanmıştır. DeneySEL sonuçlar, önerilen yaklaşımın %98.42 doğruluk oranına ulaştığını ve aynı zamanda düşük eğitim süresi ile yüksek hesaplama verimliliği sunduğunu göstermektedir. Bu bulgular, Bi-LSTM tabanlı minimalist yaklaşımların, özellikle gerçek zamanlı ve kaynak kısıtlı endüstriyel bakım uygulamaları için etkili ve uygulanabilir çözümler sunduğunu ortaya koymaktadır.

Bir diğer çalışmada, Wang ve ark. (2024), endüstriyel ortamlarda veri gizliliği kısıtları nedeniyle merkezi veri paylaşımının mümkün olmadığı durumlar için federatif öğrenme tabanlı bir rulman arıza teşhis çerçevesi önermiştir. Çalışmada, farklı istemcilerde kullanılan model mimarilerinin birbirinden farklı olması ve veri dağılımlarının bağımsız ve özdeş (non-iid) olmaması durumlarında ortaya çıkan çift yönlü heterojenite problemi ele alınmıştır. Bu kapsamda önerilen FedMAKE çerçevesi, geleneksel model ağırlığı ortalama yaklaşımlarının yetersiz kaldığı bu senaryolarda, modelden bağımsız bilgi aktarımını mümkün kılan yenilikçi bir yapı sunmaktadır. Yöntemde, değişken önemine dayalı iki adet modelden bağımsız bilgi taşıyıcı tasarlanmış ve istemciler arasında bilgi paylaşımını sağlamak amacıyla çift yönlü bilgi damıtma (bi-directional distillation) mekanizması geliştirilmiştir. Ayrıca, oluşturulan sanal örnekler aracılığıyla farklı istemciler arasında dengeli arıza bilgisi paylaşımı sağlanmıştır.

Tennessee Eastman (TE) ve CWRU veri setleri üzerinde gerçekleştirilen deneysel çalışmalar, önerilen yöntemin özellikle yüksek heterojenlik koşullarında teşhis doğruluğunu %11.7'ye kadar artırdığını göstermiştir. Bu bulgular, FedMAKE yaklaşımının federatif öğrenme tabanlı arıza teşhisi uygulamalarında etkili ve uygulanabilir bir çözüm sunduğunu ortaya koymaktadır.

Liu ve Li (2025), değişken çalışma koşullarının etkili olduğu rüzgar türbini dişli kutularında arıza teşhisi problemini ele alarak, zaman-frekans tabanlı bir derin öğrenme yaklaşımı olan time-frequency ensemble network (TFEN) modelini önermiştir. Bu modelde, sürekli dalgacık dönüşümü (CWT) kullanılarak elde edilen zaman-frekans spektrumları farklı görüntü temsilleri şeklinde işlenmiş ve bu temsiller, evrimsel sinir ağları (CNN) ile kanal dikkat mekanizmaları aracılığıyla birleştirilmiştir. Böylece, farklı arıza türleri arasındaki ayırt edici özelliklerin daha etkin bir şekilde öğrenilmesi sağlanmıştır. Önerilen yaklaşım yalnızca yüksek sınıflandırma doğruluğu sağlamakla kalmamış, aynı zamanda değişken hız ve yük koşullarına karşı da güçlü bir genelleme yeteneği göstermiştir. Modelin performansı, GAN tabanlı veri artırma teknikleri ile desteklenerek daha da iyileştirilmiştir. Kaggle platformundan temin edilen rüzgar türbini veri seti üzerinde gerçekleştirilen deneysel çalışmalar sonucunda %99.51 gibi oldukça yüksek bir doğruluk oranına ulaşılmış ve bu performansın mevcut yöntemlere kıyasla üstün olduğu ifade edilmiştir. Elde edilen bulgular, TFEN modelinin özellikle değişken çalışma koşullarına sahip endüstriyel sistemlerde güvenilir ve etkili bir arıza teşhis aracı olarak kullanılabileceğini göstermektedir.

Yu ve ark. (2025), döner makine arıza teşhisi problemlerinde etiketli veri ihtiyacını azaltmaya yönelik olarak, yarı denetimli alan uyarlama temelli CSS-DADBN (Chaotic Sparrow Search optimized Domain Adaptive Deep Belief Network) modelini geliştirmiştir. Önerilen yöntemde, etiketli kaynak alan verileri ile etiketsiz hedef alan verileri birlikte değerlendirilmiş ve hedef alan için sahte (pseudo) etiketler oluşturulmuştur. Bu etiketlerin güvenilirliğini artırmak amacıyla, kaotik serçe arama algoritması (CSSOA) kullanılarak karar eşik değerleri optimize edilmiştir. Modelin sonraki aşamasında, derin inanç ağları (DBN) aracılığıyla hem etkili özellik çıkarımı gerçekleştirilmiş hem de kaynak ve hedef alanlar arasında özellik füzyonu sağlanmıştır. Bu sayede, farklı veri dağılımlarından kaynaklanan alan kayması (domain shift) problemi azaltılarak modelin genelleme kabiliyeti artırılmıştır. Önerilen yaklaşım, PT700 ve CWRU rulman veri setleri üzerinde test edilmiş ve elde edilen sonuçlar, BODA-DBN, PSODA-DBN ve SSADA-DBN gibi mevcut yöntemlere kıyasla daha yüksek sınıflandırma doğruluğu sağladığını ortaya koymuştur. Ayrıca, PDF, ROC-AUC ve t-SNE analizleri, modelin sahte etiket üretiminde daha yüksek doğruluk sunduğunu, alanlar arası dağılım farklılıklarını etkili biçimde azalttığını ve özellik uzayında daha belirgin sınıf ayrımı sağladığını göstermiştir. Bu yönüyle çalışma, yarı denetimli alan uyarlama yaklaşımlarının, özellikle veri etiketlemenin maliyetli ve zor olduğu endüstriyel uygulamalarda, daha esnek ve güvenilir arıza teşhis sistemlerinin geliştirilmesine önemli katkılar sunduğunu ortaya koymaktadır.

Wang ve ark. (2025), endüstriyel döner makinelerde karşılaşılan az örnekli (few-shot) arıza teşhisi problemini ele alarak, yaygın olarak kullanılan Model-Agnostic Meta-Learning (MAML) yaklaşımının sınırlılıklarını gidermeye yönelik TUBR-ML (Task-Adaptive Unbiased Regularization Meta-Learning) adlı gelişmiş bir yöntem önermiştir. Çalışmada, özellikle görevler arasındaki heterojen yapıdan kaynaklanan parametre çakışmaları ve modelin belirli görevlere aşırı uyum göstermesi sonucu ortaya çıkan görev yanlılığı problemleri üzerinde durulmuştur. Bu doğrultuda, daha etkili bir öğrenme süreci sağlamak amacıyla yapılandırılmış dropout tabanlı öznelik çıkarıcı (SDFE), görev uyumlu zayıflatma modülü ve görev tarafsız düzenleyici kayıp fonksiyonu (TURL) gibi bileşenler geliştirilmiş ve MAML'in iç-döngü ve dış-döngü güncelleme mekanizmaları yeniden düzenlenmiştir. Bu iyileştirmeler sayesinde modelin farklı görevler arasında daha dengeli ve genellenebilir temsiller öğrenmesi hedeflenmiştir. Önerilen yöntem, CWRU rulman veri seti ile birlikte rüzgar türbini ve raylı sistem veri setleri üzerinde kapsamlı deneylerle değerlendirilmiştir. Elde edilen sonuçlar, TUBR-ML yaklaşımının hem 1-shot hem de 5-shot senaryolarında sırasıyla %99.26, %99.41 ve %96.44 gibi yüksek doğruluk oranlarına ulaştığını ve GMAML, klasik MAML ve GCN-Meta gibi mevcut yöntemlere kıyasla daha üstün performans sergilediğini göstermiştir. Bu bulgular, önerilen metodun özellikle sınırlı veri koşullarında ve karmaşık çalışma ortamlarında dahi güçlü bir genelleme yeteneğine sahip olduğunu ortaya koymaktadır.

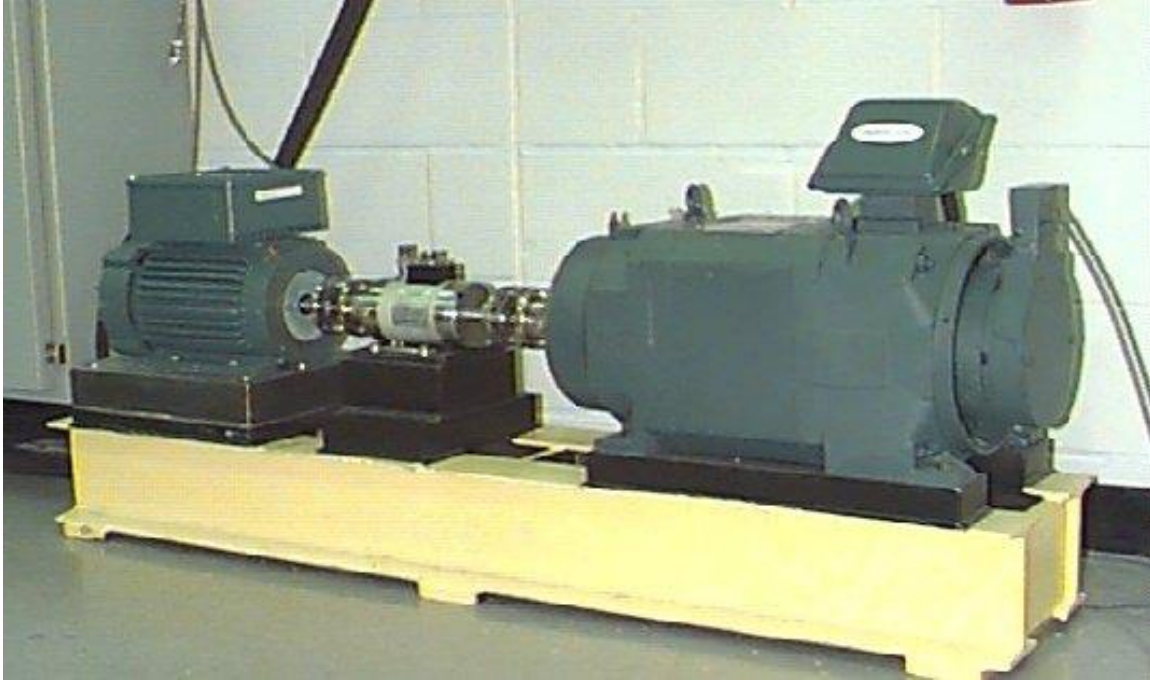
O'Neill ve ark. (2025), kriyojenik yakıt tanklarının tasarım ve işletme süreçlerinde kullanılabilecek güvenilir ısı transfer modellerinin geliştirilmesine katkı sağlamak amacıyla, sıvı azot, argon ve metan gibi kriyojenik akışkanların kaynama ve söndürme (quenching) davranışlarını kapsamlı biçimde inceleyen deneysel bir çalışma gerçekleştirmiştir. Bu kapsamda, 445 adet kararlı kaynama verisi ile 5.180 adet geçici söndürme verisi elde edilerek farklı buharlaşma rejimleri—nükleat kaynama, geçiş kaynama ve film kaynama—arasındaki kritik geçiş noktaları

detaylı olarak analiz edilmiştir. Çalışmada ayrıca, tank duvar kalınlığının bu rejimler üzerindeki etkisi incelenmiş ve özellikle ince duvarlı yapıların söndürme sürecinde önemli farklılıklar gösterdiği belirlenmiştir. Bulgular, duvar kalınlığının azalmasıyla birlikte kritik ısı akısının (CHF) belirgin şekilde düştüğünü, buna karşılık Leidenfrost noktası (minimum ısı akısı, MHF) değerinin daha düşük ısı akılarında ortaya çıktığını göstermektedir. Elde edilen geniş kapsamlı deneysel veri seti, NASA'nın kriyojenik yakıt yönetimi çalışmalarında kullanılabilecek güvenilir korelasyonların geliştirilmesine zemin hazırlamaktadır. Ayrıca sonuçlar, özellikle film kaynama bölgesinin düşük eğimli kısmının doğru modellenmesinin, uzun süreli tank soğutma performansı açısından kritik bir öneme sahip olduğunu ortaya koymaktadır.

Sonuç olarak, literatürde yer alan çalışmalar incelendiğinde, rulman arıza teşhisi alanında derin öğrenme tabanlı yaklaşımların önemli başarılar elde ettiği ve özellikle farklı çalışma koşulları, veri yetersizliği ve alan uyumsuzluğu gibi zorluklara yönelik çeşitli yenilikçi yöntemlerin geliştirildiği görülmektedir. Bununla birlikte, zamansal bağımlılıkların etkin şekilde modellenmesi ve farklı tekrarlayan sinir ağı mimarilerinin performanslarının karşılaştırmalı olarak değerlendirilmesi hâlen araştırmaya açık bir alan olarak öne çıkmaktadır. Bu bağlamda, sunulan bu çalışmada, evrişimsel sinir ağları (CNN) ile birlikte kullanılan iki farklı tekrarlayan mimari olan Uzun Kısa Süreli Bellek (LSTM) ve Geçitli Tekrarlayan Birim (GRU) modellerinin hibrit yapılar altında karşılaştırmalı analizi gerçekleştirilmiştir. Önerilen CNN-LSTM ve CNN-GRU modellerinin, hem özellik çıkarımı hem de zamansal örüntülerin öğrenilmesi açısından performansları detaylı olarak incelenmiş; elde edilen bulgular, hangi mimarinin farklı çalışma koşulları altında daha etkin sonuçlar sunduğunu ortaya koyarak literatürdeki mevcut boşluğa katkı sağlamayı amaçlamaktadır.

3.CWRU VERİ SETİ

Bu çalışmada, rulman arıza teşhisi problemini değerlendirmek amacıyla literatürde yaygın olarak kullanılan ve karşılaştırmalı analizler için bir referans niteliği taşıyan Case Western Reserve University (CWRU) rulman veri seti kullanılmıştır. CWRU veri seti, döner makinelerde arıza teşhisi çalışmalarında standart bir benchmark olarak kabul edilmekte olup, farklı arıza türleri ve çalışma koşullarını içeren kapsamlı bir veri yapısına sahiptir. Veri seti, bir elektrik motoru, tahrik sistemi ve yük mekanizmasından oluşan deney düzeneği üzerinden elde edilmiştir. Bu deney sistemi, titreşim sensörleri yardımıyla rulman davranışlarını izlemek üzere tasarlanmış olup, çalışmada kullanılan düzenek Şekil 1'de gösterilmektedir. Sistem üzerinde rulmanlara kontrollü olarak yapay hatalar uygulanmış ve bu sayede farklı arıza türlerine ait titreşim sinyalleri elde edilmiştir. Veri toplama sürecinde genellikle 12 kHz ve 48 kHz örnekleme frekansları kullanılmış ve ölçümler motorun farklı yük ve hız koşulları altında gerçekleştirilmiştir. CWRU veri seti, dört temel durumdan oluşmaktadır: normal (sağlam) durum, bilye (ball) arızası, iç bilezik (inner race) arızası ve dış bilezik (outer race) arızası. Her bir arıza türü, 0.007, 0.014 ve 0.021 inç çaplarında farklı hasar seviyeleri ile modellenmiş olup toplamda 10 farklı sınıf içermektedir (Gao ve ark., 2025).



Şekil 1. CWRU rulman test düzeneği deneysel platformu

CWRU rulman veri seti, toplamda 161 ayrı veri kümesinden oluşmakta olup, bu veri kümeleri arıza konumu ve sinyal örnekleme frekansına göre sistematik bir şekilde gruplandırılmıştır. Bu gruplar; normal (sağlam) durum, 12 kHz sürücü ucu (drive-end) arızaları, 48 kHz sürücü ucu arızaları ve 12 kHz fan ucu (fan-end) arızaları olmak üzere farklı çalışma senaryolarını kapsamaktadır. Her bir grup içerisinde üç temel arıza türü yer almakta olup bunlar iç bilezik arızası (Inner Fault, IF), bilye/yuvarlanma elemanı arızası (Ball Fault, BF) ve dış bilezik arızası (Outer Fault, OF) şeklindedir. Titreşim verileri, farklı çalışma koşullarını temsil edecek şekilde elde edilmiş olup motor hızları yaklaşık 1720 ile 1797 devir/dakika (RPM) arasında değişmekte ve sistem 0, 1, 2 ve 3 beygir gücü (HP) yük seviyelerinde test edilmiştir. Ayrıca, deney düzeneği üzerinde farklı sensör yerleşimleri kullanılarak ölçümler gerçekleştirilmiştir. Bu yerleşimler; merkez (6:00 yönü), dik/yan konum (3:00 yönü) ve karşı konum (12:00 yönü) olarak belirlenmiş olup, arızaların titreşim sinyalleri üzerindeki etkilerinin farklı açılardan gözlemlenmesine imkân tanımaktadır. Bu kapsamlı veri yapısı sayesinde, CWRU veri seti hem farklı arıza türlerinin hem de değişken çalışma koşullarının etkilerini analiz etmeye olanak sağlayarak, geliştirilen arıza teşhis modellerinin doğruluğu ve genelleme kabiliyetinin güvenilir bir şekilde değerlendirilmesine imkân sunmaktadır (Akcan, 2025).

CWRU veri seti, farklı rulman arızalarını kapsamlı ve sistematik bir biçimde temsil eden, iyi belgelenmiş ve yaygın olarak kullanılan bir veri kaynağıdır. Bu veri seti, sınıflandırma, arıza teşhisi ve kalan kullanım ömrü (RUL) tahmini gibi alanlarda gerçekleştirilen akademik çalışmalar için güçlü bir temel oluşturmaktadır. Yüksek örnekleme frekansına sahip olması, farklı arıza türlerini ve değişken çalışma koşullarını içermesi, veri setini hem zaman hem de frekans alanı analizleri açısından oldukça zengin kılmaktadır. Ayrıca, verilerin segmentasyon ve öznitelik çıkarımı süreçlerine uygun biçimde düzenlenmiş olması, araştırmacılara önemli ölçüde uygulama kolaylığı sağlamaktadır.

Bu çalışma kapsamında, CWRU veri setinden elde edilen titreşim sinyalleri kullanılarak Evrişimsel Sinir Ağı–Uzun Kısa Süreli Bellek (CNN-LSTM) ve Evrişimsel Sinir Ağı–Kapılı Tekrarlayan Birim (CNN-GRU) tabanlı hibrit derin öğrenme modelleri geliştirilmiştir. Ham titreşim sinyalleri, 2048 örnek uzunluğunda ve %50 örtüşme oranına sahip olacak şekilde segmentlere ayrılmıştır. Elde edilen her bir segment, model performansını artırmak ve veri dağılımını standardize etmek amacıyla z-puanı (z-skor) normalizasyonu uygulanarak ölçeklendirilmiştir. Sınıf etiketleri, Etiket Kodlayıcı (LabelEncoder) ve tek-sıcak kodlama (one-hot encoding) yöntemleri kullanılarak sayısal forma dönüştürülmüş; veri seti, sınıf dağılımı korunacak şekilde katmanlı (stratified) olarak %80 eğitim ve %20 test verisi olacak biçimde ayrılmıştır. Önerilen hibrit modellerde, çok katmanlı evrişimsel sinir ağı yapıları aracılığıyla anlamlı özellikler çıkarıldıktan sonra, zamansal bağımlılıkların öğrenilmesi amacıyla sırasıyla LSTM ve GRU katmanları kullanılmıştır. Model eğitim sürecinde, ağırlıkların güncellenmesi için Uyarlamalı Moment Tahmini (Adam) optimizasyon algoritması tercih edilmiştir. Bununla birlikte, aşırı öğrenme (overfitting) problemini önlemek ve

modelin genelleme yeteneğini artırmak amacıyla erken durdurma (early stopping) ve öğrenme oranı azaltma (learning rate reduction) gibi düzenleme teknikleri uygulanmıştır. Bu sayede, geliştirilen modellerin daha kararlı ve yüksek performanslı sonuçlar üretmesi sağlanmıştır.

4.METHOD

A. EVRİŞİMSSEL SİNİR AĞLARI (CNN)

Ham verilerden otomatik olarak anlamlı özellikler çıkarabilme yetenekleri sayesinde arıza teşhisi ve sınıflandırma problemlerinde yaygın olarak kullanılan derin öğrenme modellerinden biridir. Geleneksel makine öğrenmesi yaklaşımlarında özellik çıkarımı genellikle uzman bilgisi gerektirirken, CNN yapıları bu süreci otomatik hale getirerek modelin doğrudan veriden öğrenmesini sağlamaktadır (Ayadi ve ark., 2021). Bu durum, özellikle karmaşık ve doğrusal olmayan yapıya sahip titreşim sinyallerinin analizinde önemli bir avantaj sunmaktadır.

CNN mimarisi temel olarak evrişim, aktivasyon, havuzlama ve tam bağlantılı katmanlardan oluşmaktadır. Bu yapı içerisinde en kritik bileşen olan evrişim katmanı, giriş verisi üzerinde kayan filtreler aracılığıyla yerel özellikleri çıkararak özellik haritaları üretir. Bu süreçte ilk katmanlar düşük seviyeli özellikleri öğrenirken, daha derin katmanlar karmaşık ve soyut özellikleri temsil edebilmektedir (Lei ve ark., 2019). Evrişim katmanlarını takiben kullanılan doğrusal olmayan aktivasyon fonksiyonları, modele esneklik kazandırarak daha karmaşık örüntülerin öğrenilmesini mümkün kılmaktadır.

Havuzlama katmanları, özellik haritalarının boyutunu azaltarak hesaplama maliyetini düşürmekte ve modelin aşırı öğrenmesini engellemektedir. Özellikle maksimum havuzlama yöntemi, en belirgin özelliklerin korunmasını sağlayarak modelin genelleme performansını artırmaktadır. Son aşamada yer alan tam bağlantılı katmanlar ise çıkarılan özellikleri kullanarak sınıflandırma işlemini gerçekleştirmekte ve genellikle Softmax fonksiyonu ile sınıf olasılıklarını hesaplamaktadır (Sharma & Phonsa, 2021). Bu bütünsel yapı sayesinde CNN modelleri, uçtan uca öğrenme yaklaşımı ile yüksek doğrulukta sınıflandırma performansı sunabilmektedir.

B. UZUN KISA SÜRELİ BELLEK (LSTM)

Ağları, zaman serisi verilerinde uzun vadeli bağımlılıkları öğrenebilmek amacıyla geliştirilmiş özel bir tekrarlayan sinir ağı (RNN) türüdür. Geleneksel RNN yapılarında karşılaşılan gradyan kaybolması ve patlaması problemleri, LSTM mimarisinde yer alan bellek hücreleri ve kapı mekanizmaları sayesinde büyük ölçüde giderilmektedir. Bu sayede LSTM modelleri, geçmiş zaman adımlarına ait bilgileri uzun süre koruyabilmekte ve özellikle zamana bağlı değişim gösteren karmaşık sistemlerin modellenmesinde yüksek başarı sağlamaktadır (González-Planet & Juez, 2025).

LSTM mimarisi, temel olarak giriş (input), unutma (forget) ve çıkış (output) kapılarından oluşan bir yapı üzerine kuruludur. Bu kapılar, hücre durumunun (cell state) hangi bilgileri saklayacağına, güncelleyeceğine veya sileceğine karar vererek modelin önemli bilgileri filtrelemesini sağlar. Bu mekanizma sayesinde LSTM, yalnızca kısa süreli değil aynı zamanda uzun dönemli bağımlılıkları da etkin bir şekilde öğrenebilmektedir. Bu özellik, özellikle titreşim sinyalleri gibi zamansal örüntüler içeren verilerin analizinde kritik bir avantaj sunmaktadır. LSTM ağları, doğrusal olmayan ve karmaşık ilişkileri modelleme yetenekleri sayesinde enerji tüketimi tahmini, hidrolik modelleme ve yapısal sağlık izleme gibi birçok farklı alanda başarılı bir şekilde uygulanmaktadır. Özellikle çok değişkenli ve yüksek boyutlu zaman serisi verilerinde, hem içsel dinamikleri hem de dışsal faktörleri dikkate alarak güçlü tahmin performansı sunmaktadır (Pai ve ark., 2025). Ayrıca, CNN ile birlikte hibrit olarak kullanıldığında, CNN tarafından çıkarılan uzaysal özelliklerin LSTM aracılığıyla zamansal bağımlılıklarla ilişkilendirilmesi mümkün olmakta ve bu sayede modelin genel performansı önemli ölçüde artırılmaktadır (Minh ve ark., 2025).

C. KAPILI TEKRARLAYAN BİRİM (GRU)

Zaman serisi verilerinde uzun dönemli bağımlılıkların öğrenilmesini amaçlayan ve tekrarlayan sinir ağlarının (RNN) geliştirilmiş bir türevi olarak önerilen bir derin öğrenme mimarisidir. Geleneksel RNN yapılarında ortaya çıkan gradyan kaybolması ve patlaması problemlerini azaltmak amacıyla geliştirilen GRU, kapı mekanizması sayesinde bilgi akışını kontrol ederek daha stabil ve etkili bir öğrenme süreci sunmaktadır. Bu yapı, özellikle ardışık verilerde

geçmiş bilginin ne kadarının korunacağına ve ne kadarının güncelleneceğine adaptif olarak karar verebilmekte, böylece zaman bağımlı karmaşık örüntülerin modellenmesinde yüksek performans sağlamaktadır. GRU mimarisi, LSTM'ye kıyasla daha sade bir yapıya sahip olup iki temel kapıdan oluşmaktadır: güncelleme kapısı (update gate) ve sıfırlama kapısı (reset gate). Güncelleme kapısı, önceki gizli durum bilgisinin ne kadarının korunacağını belirlerken; sıfırlama kapısı, geçmiş bilginin mevcut hesaplamada ne ölçüde dikkate alınacağını kontrol etmektedir. Bu iki kapı mekanizması sayesinde GRU, hem kısa hem de uzun vadeli bağımlılıkları etkin bir şekilde öğrenebilmekte ve hesaplama maliyetini azaltarak daha hızlı eğitim süreci sunmaktadır (Zhang ve ark., 2025). Ayrıca, GRU'nun LSTM'ye göre daha az parametre içermesi, modelin daha düşük hesaplama yükü ile benzer performans göstermesine olanak tanımaktadır. Matematiksel olarak GRU, güncelleme ve sıfırlama kapıları aracılığıyla aday gizli durumu hesaplamakta ve önceki gizli durum ile bu aday durumu ağırlıklı olarak birleştirerek nihai gizli durumu oluşturmaktadır. Bu süreç, modelin geçmiş ve mevcut bilgi arasında dengeli bir temsil oluşturmasını sağlamaktadır (Yunita ve ark., 2025). Bu özellikleri sayesinde GRU, özellikle sensör verileri, titreşim sinyalleri ve sağlık izleme gibi zaman serisi tabanlı uygulamalarda yaygın olarak kullanılmakta ve yüksek doğruluk oranları ile dikkat çekmektedir.

D. PERFORMANS METRİKLERİ

Makine öğrenmesi ve yapay zekâ tabanlı sınıflandırma modellerinin başarısını değerlendirmek amacıyla kullanılan performans metrikleri, modelin tahmin yeteneğini nicel olarak ortaya koyan temel ölçütlerdir. Bu metrikler, modelin yalnızca genel doğruluğunu değil, aynı zamanda farklı sınıfları ne derece etkin ayırt edebildiğini de analiz etmeye imkân tanımaktadır. Literatürde en yaygın kullanılan performans ölçütleri arasında doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1-skoru (F1-score) yer almakta olup, her biri sınıflandırma performansının farklı bir yönünü temsil etmektedir.

Doğruluk (Accuracy), model tarafından yapılan tüm tahminler içerisinde doğru sınıflandırılan örneklerin oranını ifade etmektedir. Bu metrik, genel performans hakkında hızlı bir değerlendirme sunsa da özellikle sınıf dağılımının dengesiz olduğu durumlarda yanıltıcı olabilir. Doğruluk aşağıdaki şekilde hesaplanmaktadır:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Burada TP doğru pozitif, TN doğru negatif, FP yanlış pozitif ve FN yanlış negatif değerlerini ifade etmektedir.

Kesinlik (Precision), modelin pozitif olarak sınıflandırdığı örneklerin ne kadarının gerçekten pozitif olduğunu göstermektedir. Özellikle yanlış pozitif sonuçların maliyetli olduğu uygulamalarda önemli bir performans ölçütüdür. Precision aşağıdaki şekilde tanımlanmaktadır:

$$precision = \frac{TP}{TP + FP}$$

Yüksek precision değeri, modelin yanlış alarm üretme oranının düşük olduğunu göstermektedir.

Duyarlılık (Recall), gerçek pozitif örneklerin model tarafından doğru şekilde tespit edilme oranını ifade etmektedir. Bu metrik, özellikle yanlış negatif sonuçların kritik olduğu sistemlerde büyük önem taşımaktadır. Recall değeri şu şekilde hesaplanmaktadır:

$$recall = \frac{TP}{TP + FN}$$

Yüksek recall değeri, modelin pozitif örnekleri kaçırmadan tespit edebildiğini göstermektedir.

F1-Skoru (F1-Score) ise precision ve recall metriklerinin harmonik ortalaması olup, bu iki ölçüt arasında dengeli bir performans değerlendirmesi sunmaktadır. Özellikle veri setlerinin dengesiz olduğu durumlarda daha güvenilir bir ölçüt olarak kabul edilmektedir. F1-skoru aşağıdaki şekilde ifade edilmektedir:

$$f1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

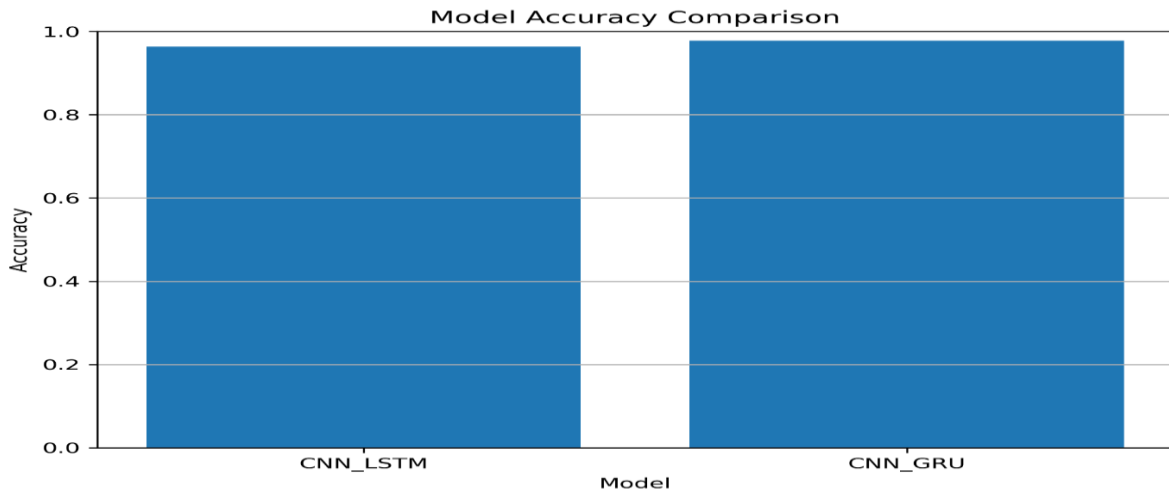
Sonuç olarak, bu metriklerin birlikte değerlendirilmesi, bir sınıflandırma modelinin hem genel performansını hem de hata türlerine karşı davranışını daha kapsamlı ve güvenilir bir şekilde analiz etmeye olanak sağlamaktadır (Akcan, 2025).

5.BULGULAR

Bu bölümde, rulman arıza teşhisi için önerilen derin öğrenme tabanlı modellerin performansları kapsamlı bir şekilde değerlendirilmiştir. Case Western Reserve University (CWRU) rulman veri setinden elde edilen titreşim sinyalleri kullanılarak geliştirilen **CNN-LSTM ve CNN-GRU hibrit modelleri uygulanmış** ve her iki modelin performansı doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1-skoru metrikleri açısından detaylı olarak analiz edilmiştir. Ayrıca, modellerin eğitim süreci; kayıp (loss) ve doğruluk (accuracy) eğrileri üzerinden incelenmiş, elde edilen sonuçlar grafikler ve karışıklık matrisleri ile görselleştirilerek karşılaştırmalı olarak değerlendirilmiştir. Performansın yalnızca genel doğruluk düzeyinde değil, aynı zamanda sınıf bazlı analizler üzerinden incelenmesi, hangi arıza türlerinin daha yüksek doğrulukla tespit edildiğini ve hangi sınıflarda hata oranlarının arttığını ortaya koyması açısından büyük önem taşımaktadır. Bu kapsamda, önerilen iki modelin sınıflandırma başarıları ve karşılaştırmalı performans sonuçları tablo 1 ve Şekil 2’de sunulmuştur.

Tablo 1. Önerilen hibrit modellerin test doğruluk performansları

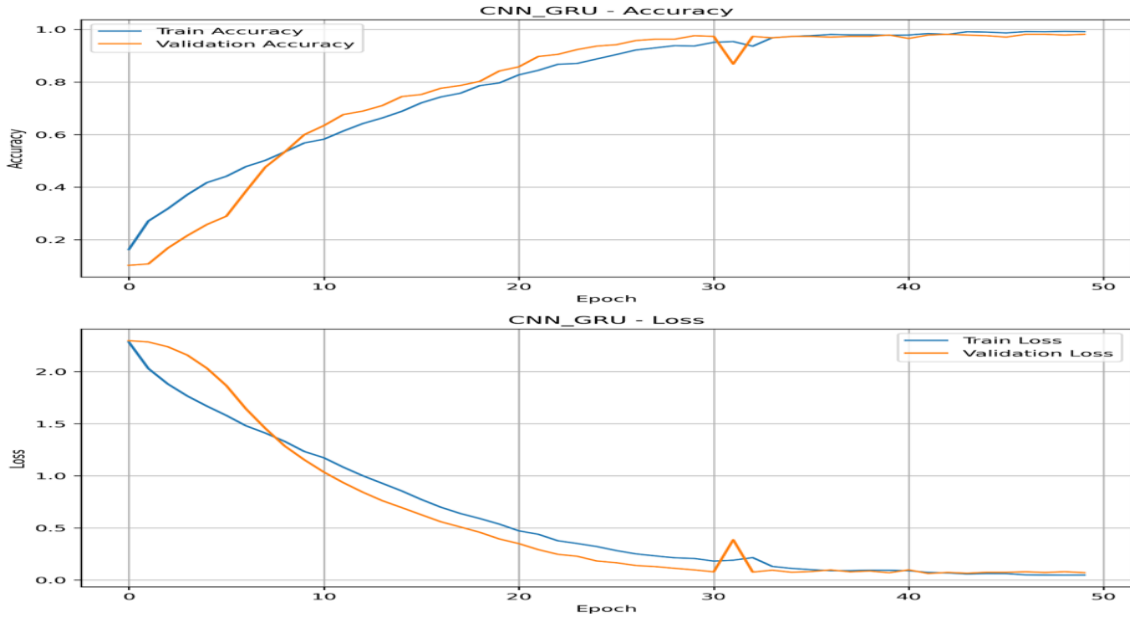
Model	Test_Accuracy
CNN_LSTM	0.9632
CNN_GRU	0.9779



Şekil 2. CNN-LSTM ve CNN-GRU modellerinin doğruluk (accuracy) performanslarının karşılaştırılması

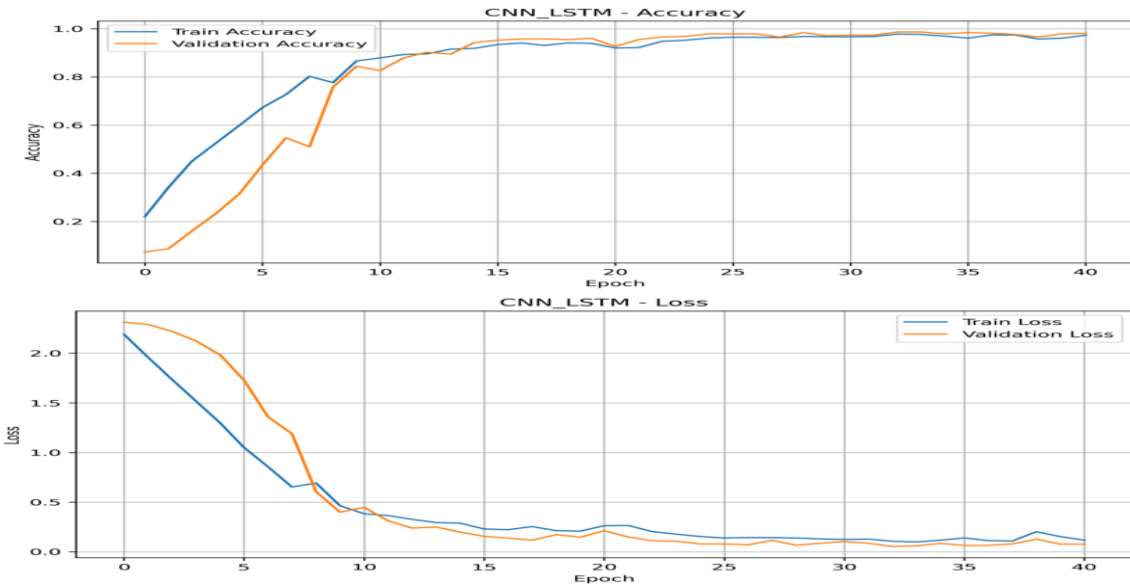
Şekil 3’te, CNN-GRU modelinin eğitim sürecine ait doğruluk (accuracy) ve kayıp (loss) değişimleri epoch’a bağlı olarak gösterilmektedir. Üst grafikte eğitim ve doğrulama doğruluk değerlerinin epoch ilerledikçe düzenli bir şekilde arttığı ve yaklaşık %98–99 seviyelerine ulaştığı görülmektedir. Eğitim ve doğrulama eğrilerinin birbirine oldukça yakın seyretmesi, modelin iyi bir genelleme performansı sergilediğini ve aşırı öğrenme (overfitting) probleminin oluşmadığını göstermektedir. Alt grafikte ise hem eğitim hem de doğrulama kayıp değerlerinin epoch boyunca istikrarlı bir şekilde azaldığı gözlemlenmektedir. Yaklaşık 30. epoch civarında kısa süreli bir dalgalanma görülsede

modelin genel olarak kararlı bir öğrenme süreci gerçekleştirdiği anlaşılmaktadır. Bu sonuçlar, CNN-GRU modelinin hızlı ve stabil bir şekilde öğrenerek yüksek performans elde ettiğini ortaya koymaktadır.



Şekil 3. CNN-GRU modeline ait eğitim ve doğrulama doğruluk (accuracy) ile kayıp (loss) eğrileri

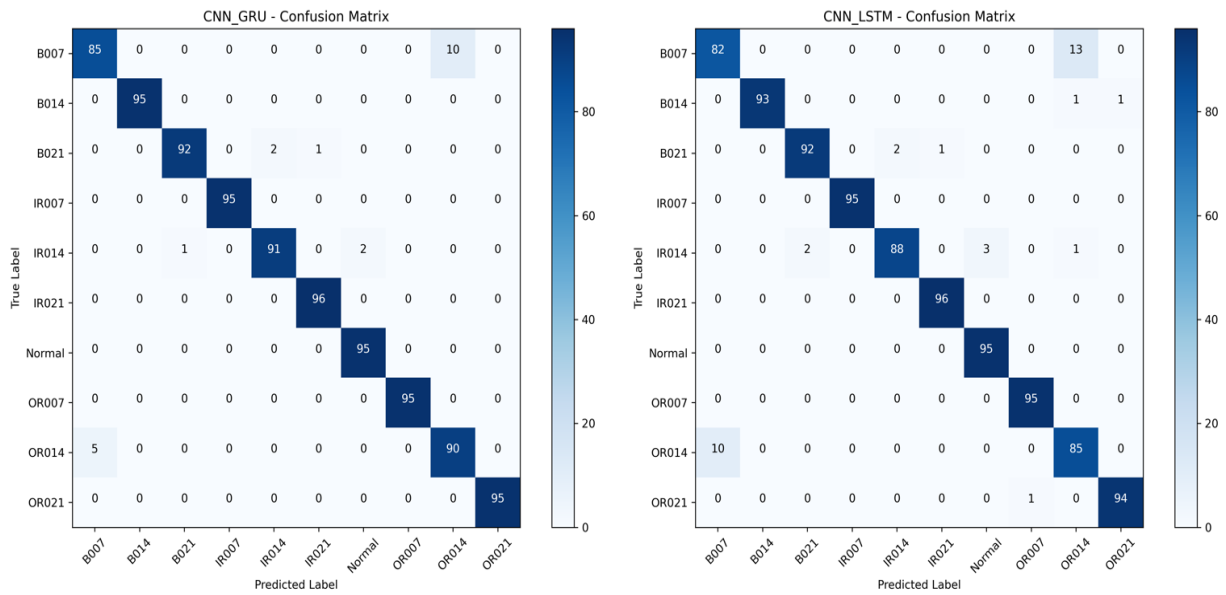
Şekil 4'te, CNN-LSTM modelinin eğitim sürecine ait doğruluk (accuracy) ve kayıp (loss) değerlerinin epoch'a bağlı değişimi gösterilmektedir. Üst grafikte eğitim ve doğrulama doğruluk değerlerinin epoch ilerledikçe hızlı bir şekilde arttığı ve yaklaşık %97–99 seviyelerine ulaştığı görülmektedir. Özellikle 10. epoch sonrasında doğrulama doğruluğunda belirgin bir artış yaşandığı ve eğitim doğruluğu ile oldukça yakın seyrettiği dikkat çekmektedir. Bu durum, modelin öğrenme sürecinin başarılı olduğunu ve genelleme kabiliyetinin yüksek olduğunu göstermektedir. Alt grafikte ise eğitim ve doğrulama kayıp değerlerinin epoch boyunca düzenli olarak azaldığı gözlemlenmektedir. Başlangıçta yüksek olan kayıp değerleri, ilerleyen epoch'larda önemli ölçüde düşerek düşük seviyelerde stabilize olmuştur. Bununla birlikte, bazı epoch'larda küçük dalgalanmalar görülmekle birlikte genel eğilim istikrarlı bir azalış yönündedir. Eğitim ve doğrulama kayıp eğrilerinin birbirine yakın seyretmesi, modelin aşırı öğrenme (overfitting) problemi yaşamadığını göstermektedir. Genel olarak değerlendirildiğinde, CNN-LSTM modelinin kararlı bir öğrenme süreci gerçekleştirdiği ve yüksek doğruluk değerlerine ulaşarak rulman arıza teşhisi probleminde etkili bir performans sergilediği söylenebilir.



Şekil 4. CNN-LSTM modeline ait eğitim ve doğrulama doğruluk (accuracy) ile kayıp (loss) eğrileri

Şekil 5'te, CNN-GRU ve CNN-LSTM modellerine ait sınıflandırma performanslarını gösteren karışıklık matrisleri sunulmaktadır. Her iki modelde de diyagonal elemanların yüksek olması, sınıfların büyük oranda doğru şekilde sınıflandırıldığını göstermektedir. Ancak modeller arasında bazı sınıflar açısından dikkat çekici farklılıklar bulunmaktadır. CNN-GRU modelinde, özellikle B014, IR007, IR021, Normal, OR007 ve OR021 sınıflarında neredeyse hatasız sınıflandırma elde edilmiştir. Bununla birlikte, B007 ve OR014 sınıflarında sınırlı sayıda yanlış sınıflandırma gözlemlenmiştir. Örneğin, B007 sınıfına ait bazı örneklerin OR014 olarak tahmin edildiği görülmektedir. CNN-LSTM modelinde ise genel doğruluk yüksek olmakla birlikte, bazı sınıflarda hata oranlarının arttığı dikkat çekmektedir. Özellikle OR014 sınıfında daha fazla yanlış sınıflandırma (yaklaşık 10 örnek) meydana gelmiş ve bu sınıfın diğer sınıflarla karıştırılma eğilimi gösterdiği görülmüştür. Ayrıca B007 ve IR014 sınıflarında da CNN-GRU modeline kıyasla daha fazla hata yapıldığı gözlemlenmiştir.

Genel olarak değerlendirildiğinde, her iki model de yüksek sınıflandırma performansı sergilemekle birlikte, CNN-GRU modelinin sınıflar arası ayrımı daha başarılı gerçekleştirdiği ve özellikle zor ayrıştırılan sınıflarda daha düşük hata oranına sahip olduğu görülmektedir. Bu sonuçlar, CNN-GRU modelinin rulman arıza teşhisi probleminde daha güvenilir bir performans sunduğunu ortaya koymaktadır.



Şekil 5. CNN-GRU ve CNN-LSTM modellerine ait karışıklık matrislerinin karşılaştırılması

Tablo 2'de, CNN-GRU modelinin her bir arıza sınıfı için elde ettiği kesinlik (precision), duyarlılık (recall) ve F1-skoru değerleri sunulmaktadır. Sonuçlar incelendiğinde, modelin çoğu sınıfta oldukça yüksek performans sergilediği ve birçok sınıf için bu metriklerin 0.98–1.00 aralığında olduğu görülmektedir. Özellikle B014, IR007, OR007 ve OR021 sınıflarında tüm metriklerin 1.00 değerine ulaşması, modelin bu sınıfları hatasız şekilde ayırt edebildiğini göstermektedir. Bununla birlikte, bazı sınıflarda görece daha düşük performans değerleri dikkat çekmektedir. Özellikle B007 sınıfında recall değerinin 0.89 olması, bu sınıfa ait bazı örneklerin model tarafından doğru şekilde tespit edilemediğini göstermektedir. Benzer şekilde, OR014 sınıfında precision (0.90) ve F1-skoru (0.92) değerlerinin diğer sınıflara göre daha düşük olması, bu sınıfın diğer arıza türleri ile karıştırılma eğilimi gösterdiğini ortaya koymaktadır.

Genel olarak değerlendirildiğinde, CNN-GRU modelinin sınıf bazlı performansının oldukça yüksek olduğu ve özellikle çoğu arıza türünde güçlü ayırt edici özellikler öğrenebildiği görülmektedir. Bununla birlikte, sınıflar arası benzerliğin yüksek olduğu bazı durumlarda performansın kısmen düştüğü, ancak genel başarının yüksek seviyede korunduğu anlaşılmaktadır.

Tablo 2. CNN-GRU modeline ait sınıf bazlı performans metrikleri

Sınıf	Precision	Recall	F1-Score
B007	0,89	0,86	0,88
B014	1,00	0,98	0,99
B021	0,98	0,97	0,97
IR007	1,00	1,00	1,00
IR014	0,98	0,94	0,96
IR021	0,99	1,00	0,99
Normal	0,97	1,00	0,98
OR007	0,99	1,00	0,99
OR014	0,85	0,89	0,87
OR021	0,99	0,99	0,99

Tablo 3'te, CNN-LSTM modelinin her bir arıza sınıfı için elde ettiği kesinlik (precision), duyarlılık (recall) ve F1-skoru değerleri sunulmaktadır. Sonuçlar incelendiğinde, modelin genel olarak yüksek performans sergilediği ve birçok sınıfta metrik değerlerinin 0.96–1.00 aralığında olduğu görülmektedir. Özellikle B014, IR007, IR021, OR007 ve OR021 sınıflarında oldukça yüksek doğruluk elde edilmiş, bazı sınıflarda ise metriklerin 1.00 seviyesine ulaştığı gözlemlenmiştir. Bununla birlikte, bazı sınıflarda performansın görece daha düşük olduğu dikkat çekmektedir. Özellikle B007 sınıfında recall değerinin 0.86 olması, bu sınıfa ait örneklerin bir kısmının model tarafından doğru şekilde tespit edilemediğini göstermektedir. Ayrıca OR014 sınıfında precision (0.85) ve F1-skoru (0.87) değerlerinin diğer sınıflara kıyasla daha düşük olması, bu sınıfın diğer arıza türleri ile karıştırılma eğiliminin daha yüksek olduğunu ortaya koymaktadır. Bunun yanı sıra, IR014 sınıfında recall değerinin 0.94 olması, bu sınıf için de sınırlı düzeyde hata yapıldığını göstermektedir.

Genel olarak değerlendirildiğinde, CNN-LSTM modelinin sınıf bazlı performansının yüksek olduğu ancak bazı zor ayrıştırılan sınıflarda CNN-GRU modeline kıyasla daha fazla hata yaptığı söylenebilir. Buna rağmen modelin genel doğruluğunun yüksek olması, rulman arıza teşhisi probleminde etkin bir şekilde kullanılabileceğini göstermektedir.

Tablo 3. CNN-LSTM modeline ait sınıf bazlı performans metrikleri

Sınıf	Precision	Recall	F1-Score
B007	0,94	0,89	0,92
B014	1,00	1,00	1,00
B021	0,99	0,97	0,98
IR007	1,00	1,00	1,00
IR014	0,98	0,97	0,97
IR021	0,99	1,00	0,99
Normal	0,98	1,00	0,99
OR007	1,00	1,00	1,00
OR014	0,90	0,95	0,92
OR021	1,00	1,00	1,00

Elde edilen bulgular genel olarak değerlendirildiğinde, hem CNN-LSTM hem de CNN-GRU hibrit derin öğrenme modellerinin rulman arıza teşhisi probleminde yüksek doğruluk oranları ile başarılı sonuçlar ürettiği görülmektedir. CNN-LSTM modeli %96.32 test doğruluğu elde ederken, CNN-GRU modeli %97.79 test doğruluğuna ulaşarak daha yüksek bir performans sergilemiştir. Her iki model de eğitim sürecinde kararlı bir öğrenme davranışı göstermiş, düşük kayıp değerlerine ulaşmış ve yüksek doğruluk seviyelerinde stabilize olmuştur. Bununla birlikte, sınıf bazlı analizler incelendiğinde, özellikle zor ayrıştırılan sınıflarda (B007 ve OR014 gibi) CNN-GRU modelinin daha düşük hata oranı ile daha başarılı sonuçlar ürettiği dikkat çekmektedir. CNN-LSTM modeli de güçlü bir performans ortaya koymasına rağmen, bazı sınıflarda daha fazla karışıklık yaşadığı gözlemlenmiştir. Sonuç olarak, GRU tabanlı yapının daha az parametre ile daha hızlı ve etkili öğrenme gerçekleştirmesi, CNN-GRU modelinin bu çalışmada daha yüksek

doğruluk ve daha iyi genelleme performansı elde etmesinin temel nedenlerinden biri olarak değerlendirilmektedir. Bu bulgular, rulman arıza teşhisi uygulamalarında CNN-GRU tabanlı hibrit modellerin daha avantajlı bir yaklaşım sunduğunu ortaya koymaktadır.

6.TARTIŞMA VE SONUÇ

Bu çalışmada, rulman arıza teşhisi problemi için derin öğrenme tabanlı hibrit yaklaşımlar geliştirilmiş ve CNN-LSTM ile CNN-GRU modelleri karşılaştırmalı olarak analiz edilmiştir. Case Western Reserve University (CWRU) veri seti kullanılarak gerçekleştirilen deneysel çalışmalar sonucunda her iki modelin de yüksek doğruluk oranları ile başarılı sonuçlar ürettiği belirlenmiştir. Ancak elde edilen bulgular, modeller arasında belirli performans farklılıklarının bulunduğunu açıkça ortaya koymaktadır.

Deneysel sonuçlara göre, CNN-LSTM modeli %96.32 test doğruluğu elde ederken, CNN-GRU modeli %97.79 test doğruluğuna ulaşarak daha üstün bir performans sergilemiştir. Bu fark her ne kadar sınırlı görünse de, sınıf bazlı analizler ve karışıklık matrisleri incelendiğinde CNN-GRU modelinin daha tutarlı ve dengeli sonuçlar ürettiği anlaşılmaktadır. Özellikle B007 ve OR014 gibi ayrıştırılması zor sınıflarda CNN-GRU modelinin daha düşük hata oranına sahip olması, bu modelin daha güçlü bir genelleme yeteneğine sahip olduğunu göstermektedir.

Eğitim süreci analizleri de bu durumu desteklemektedir. CNN-GRU modeline ait doğruluk ve kayıp eğrileri, daha stabil bir öğrenme sürecine işaret ederken, CNN-LSTM modelinde belirli epoch'larda daha fazla dalgalanma gözlemlenmiştir. Bu durum, GRU mimarisinin daha sade yapısı ve daha az parametre içermesi sayesinde daha hızlı ve etkili öğrenme gerçekleştirdiğini göstermektedir. Literatürde de GRU tabanlı modellerin, özellikle sınırlı veri ve karmaşık zaman serisi problemlerinde LSTM'ye kıyasla benzer veya daha iyi performans sağlayabildiği belirtilmektedir.

Sınıf bazlı performans metrikleri incelendiğinde, her iki modelin de birçok sınıfta precision, recall ve F1-score değerlerinin 0.95'in üzerinde olduğu görülmektedir. Bununla birlikte, CNN-LSTM modelinde bazı sınıflarda (özellikle OR014 ve B007) performans düşüşü daha belirgin iken, CNN-GRU modelinde bu düşüşün daha sınırlı kaldığı dikkat çekmektedir. Bu durum, GRU'nun zamansal bağımlılıkları daha etkin modelleyerek benzer sinyal karakteristiklerine sahip sınıflar arasındaki ayrımı daha başarılı gerçekleştirdiğini göstermektedir.

Bu çalışmanın bir diğer önemli katkısı, CNN tabanlı özellik çıkarımı ile RNN tabanlı zamansal modellemenin birlikte kullanılmasıdır. CNN katmanları sayesinde titreşim sinyallerinden ayırt edici uzaysal özellikler elde edilirken, LSTM ve GRU katmanları bu özelliklerin zaman içerisindeki değişimini modelleyerek arıza teşhis performansını artırmıştır. Elde edilen sonuçlar, hibrit derin öğrenme yaklaşımlarının geleneksel yöntemlere kıyasla daha yüksek doğruluk ve daha güçlü genelleme yeteneği sunduğunu açıkça ortaya koymaktadır.

Sonuç olarak, bu çalışmada önerilen hibrit modeller arasında CNN-GRU yaklaşımının daha yüksek doğruluk, daha stabil eğitim süreci ve daha iyi sınıf ayrımı sağlaması nedeniyle rulman arıza teşhisi problemleri için daha uygun bir yöntem olduğu belirlenmiştir. Gelecek çalışmalarda, önerilen modelin farklı veri setleri ve gerçek zamanlı endüstriyel uygulamalar üzerinde test edilmesi, modelin genelleme kabiliyetinin daha kapsamlı şekilde değerlendirilmesine katkı sağlayacaktır. Ayrıca, dikkat mekanizmaları (attention), transformer tabanlı yapılar veya çoklu sensör verilerinin entegrasyonu ile model performansının daha da artırılması mümkün olabilir.

KAYNAKÇA

Huo, C., Jiang, Q., Shen, Y., Lin, X., Zhu, Q., & Zhang, Q. (2023). A class-level matching unsupervised transfer learning network for rolling bearing fault diagnosis under various working conditions. *Applied Soft Computing*, 146, 110739.

Wu, Z., Jiang, H., Zhu, H., & Wang, X. (2023). A knowledge dynamic matching unit-guided multi-source domain adaptation network with attention mechanism for rolling bearing fault diagnosis. *Mechanical Systems and Signal Processing*, 189, 110098.

- Li, F., Wang, L., Wang, D., Wu, J., & Zhao, H. (2023). An adaptive multiscale fully convolutional network for bearing fault diagnosis under noisy environments. *Measurement*, 216, 112993.
- Zhang, J., Chen, J., Deng, H., & Hu, W. (2023). A novel framework based on adaptive Multi-Task learning for bearing fault diagnosis. *Energy Reports*, 9, 522-531.
- Wang, P., Xiong, H., & He, H. (2023). Bearing fault diagnosis under various conditions using an incremental learning-based multi-task shared classifier. *Knowledge-Based Systems*, 266, 110395.
- Hou, Y., Wang, J., Chen, Z., Ma, J., & Li, T. (2023). Diagnosisformer: An efficient rolling bearing fault diagnosis method based on improved Transformer. *Engineering Applications of Artificial Intelligence*, 124, 106507.
- Gu, X., Yu, Y., Guo, L., Gao, H., & Luo, M. (2023). CSWGAN-GP: A new method for bearing fault diagnosis under imbalanced condition. *Measurement*, 217, 113014.
- Tang, G., Yi, C., Liu, L., Yang, X., Xu, D., Zhou, Q., & Lin, J. (2023). A novel transfer learning network with adaptive input length selection and lightweight structure for bearing fault diagnosis. *Engineering Applications of Artificial Intelligence*, 123, 106395.
- Zhao, K., Xiao, J., Li, C., Xu, Z., & Yue, M. (2023). Fault diagnosis of rolling bearing using CNN and PCA fractal based feature extraction. *Measurement*, 113754.
- Yi, H., Li, D., Lu, Z., Jin, Y., Duan, H., Hou, L., Duraihem, F. Z., Awwad, E. M., & Saeed, N. A. (2025). *VibrMamba: A lightweight Mamba based fault diagnosis of rotating machinery using vibration signal*. *Measurement*, 249, 116881.
- Noussis, A., Saif, A., Khatab, A., & Diallo, C. (2024). *Using a Minimalist Bi-LSTM for Multi-Faceted Bearing Fault Detection*. *IFAC PapersOnLine*, 58(19), 1258–1263.
- Wang, J., Song, P., & Zhao, C. (2024). *Towards consensual representation: Model-agnostic knowledge extraction for dual heterogeneous federated fault diagnosis*. *Neural Networks*, 179, 106618.
- Liu, Y., & Li, W. (2025). *Time-frequency ensemble network for wind turbine gearbox fault diagnosis under variable working conditions*. *Engineering Science and Technology, an International Journal*, 38, 102410.
- Yu, W., Zhang, Q., Feng, X., Wang, K., Cai, Y., & Wu, D. (2025). *Threshold-optimized and features-fused semi-supervised domain adaptive bearing fault diagnosis via chaotic sparrow search and deep belief network*. *Neurocomputing*, 633, 128992.
- Wang, H., Lv, D., Lin, T., Han, C., & Song, L. (2025). *Task-adaptive unbiased regularization meta-learning for few-shot cross-domain fault diagnosis*. *Engineering Applications of Artificial Intelligence*, 144, 110200.
- O'Neill, L. E., Izenzon, M. G., Kharangate, C. R., Pirnstill, L., & Li, J. (2025). *Steady boiling and transient quenching of cryogenic fluids: new experiments to enable advanced propellant tank design and operations planning*. *Cryogenics*, 151, 104177.
- Gao, T., Xie, X., Wei, N., Shu, M., & Geng, Y. (2025). *Advancing few-shot rolling element bearing diagnostics with fine-grained similarity learning across variable working environments*. *Neurocomputing*, 131093.
- Akcan, E. (2025). *Detection of Bearing Faults from Vibration Signals*. *Balkan Journal of Electrical and Computer Engineering*, 13(3), 295-306.
- Ayadi, W., Elhamzi, W., Charfi, I., & Atri, M. (2021). *Deep CNN for brain tumor classification*. *Neural processing letters*, 53(1), 671-700.

- Lei, X., Pan, H., & Huang, X. (2019). A dilated CNN model for image classification. IEEE access, 7, 124087-124095.
- Sharma, A., & Phonsa, G. (2021). Image classification using CNN. In *Proceedings of the international conference on innovative computing & communication (ICICC)*.
- González-Planet, I. C., & Juez, C. (2025). Hydrological insights from a comparative evaluation of LSTM and MC-LSTM networks in the Ebro River basin (Spain). *Journal of Hydrology: Regional Studies*, 61, 102719.
- Pai, A., Mishra, K. K., Jeyan, J. M. L., & Sayal, A. (2025). Enhanced household energy consumption forecasting using multivariate long short-term memory (LSTM) networks with weather data integration. *Results in Engineering*, 106512.
- Minh, T. Q., Matos, J. C., Sousa, H. S., Ngoc, S. D., Van, T. N., Nguyen, H. X., & Nguyễn, Q. (2025). Data reconstruction leverages one-dimensional convolutional neural networks (1DCNN) combined with long short-term memory (LSTM) networks for structural health monitoring (SHM). *Measurement*, 253, 117810.
- Zhang, N., Li, J., Ma, Y., & Wu, K. (2025). Lithium-Ion Batteries state of health estimation based on optimized TCN-GRU-WNN. *Energy Reports*, 13, 2502-2515.
- Yunita, A., Pratama, M. I., Almuzakki, M. Z., Ramadhan, H., Akhir, E. A. P., Mansur, A. B. F., & Basori, A. H. (2025). Performance analysis of neural network architectures for time series forecasting: A comparative study of RNN, LSTM, GRU, and hybrid models. *MethodsX*, 15, 103462.