# Rulman Arıza Teşhisi için Derin Öğrenme: Zaman ve Frekans Alanı

*Zeliha H. Karapinar*[1], *İ. Sina Kuseyri*[2]

[1]TÜBİTAK Ulusal Metroloji Enstitüsü, Gebze Yerleşkesi-Kocaeli, Türkiye
{zeliha.karapinar}@tubitak.gov.tr

[2]Marmara Üniversitesi Makine Mühendisliği Bölümü, Maltepe-İstanbul, Türkiye
{sina.kuseyri}@marmara.edu.tr

## Özetçe

Rulman arızalarının erken tespiti, endüstriyel makinelerin güvenilirliğini, maliyet verimliliğini ve operasyonel sürekliliğini artırmada çok önemli bir rol oynar. Geleneksel olarak, belirli arıza modelleri dönüşümden sonra daha ayırt edilebilir hale geldiği için, rulman arıza teşhisinde frekans alanı analizi mühendislik uygulamalarında tercih edilmiştir. Yaygın bir varsayım, bu dönüşüm sırasında faz gibi ince zaman alanı bilgilerinin yitirilmesinin teşhis başarımını düşürdüğü yönündedir; ancak bu varsayım sistematik biçimde sınanmamıştır. Bu çalışmada, ham zaman serisi titreşim verileri ve Ayrık Fourier Dönüşümü (DFT) tabanlı frekans alanı genlik spektrumu üzerinde ayrı ayrı eğitilmiş tek boyutlu evrişimsel sinir ağlarının (1D-CNN'ler) sızıntısız ve karşılaştırmalı bir değerlendirmesi sunulmuştur. Modeller yalnızca yük seviyeleri 0–2 ile eğitilmiş, tamamen görülmemiş yük seviyesi 3 kayıtlarında test edilmiş ve her yapılandırma üç farklı rastgele tohum ile tekrarlanmıştır. Bulgular, görülmemiş çalışma koşullarında DFT tabanlı frekans alanı temsilinin ham zaman alanı temsiline kıyasla belirgin biçimde daha iyi genelleştiğini ortaya koymaktadır. Bu sonuç, faz bilgisinin dışlanmasının tek başına teşhis performansında bir kayba yol açmadığını ve bir temsilin belirli bilgileri kaybetmesinin, görev açısından anlamlı bilgilerin de zorunlu olarak kaybedildiği anlamına gelmediğini göstermektedir. Dolayısıyla, zaman ve frekans alanı temsilleri arasındaki fark yalnızca bilgi korunumu açısından değil, görev açısından anlamlı temsil ve farklı çalışma koşullarına genelleme açısından da değerlendirilmelidir. Elde edilen performans farkı ayrıca, her iki temsil kullanılarak elde edilen açıklamaların modele ne ölçüde sadık olduğu ve rulman arızalarının fiziksel karakteristikleriyle ne ölçüde uyum gösterdiğinin incelenmesi için bir temel oluşturmaktadır.

## Abstract

Early detection of bearing faults plays a crucial role in enhancing the reliability, cost efficiency, and operational continuity of industrial machinery. Traditionally, frequency-domain analysis has been favored in engineering practice for bearing fault diagnosis, since certain fault patterns become more distinguishable after transformation. A common assumption is that discarding subtle time-domain information such as phase during this transformation degrades diagnostic performance; however, this assumption has not been tested systematically. In this study, a leakage-free comparative evaluation of one-dimensional convolutional neural networks (1D-CNNs) trained separately on raw time-series vibration data and on Discrete Fourier Transform (DFT)-based frequency-domain magnitude spectra was presented. The models were trained only on load levels 0–2, tested on entirely unseen load level 3 recordings, and each configuration was repeated with three different random seeds. The findings revealed that, under unseen operating conditions, the DFT-based frequency-domain representation generalized markedly better than the raw time-domain representation. This result showed that discarding phase information does not by itself cause a loss in diagnostic performance, and that a representation losing certain information does not necessarily mean that task-relevant information is also lost. Consequently, the difference between time- and frequency-domain representations should be assessed not only in terms of information preservation, but also in terms of task-relevant representation and generalization to different operating conditions. The observed performance difference also provides a basis for investigating the extent to which the explanations obtained from both representations are faithful to the model and consistent with the physical characteristics of bearing faults.

## 1. Giriş

Rulmanlar, dönen makinelerde temel bileşenlerdir ve hareketli parçalar arasındaki sürtünmeyi azaltarak düzgün hareketi sağlarlar. Bununla birlikte, değişken yükler, zorlu çalışma koşulları ve uzun süreli kullanım nedeniyle aşamalı aşınmaya ve beklenmedik arızalara eğilimlidirler [1]. Bu tür öngörülemeyen arızalar, planlanmamış duruş süreleri, maliyetli onarımlar ve potansiyel güvenlik tehlikeleri nedeniyle önemli ekonomik kayıplara yol açabilecek felaket niteliğinde ekipman arızalarına neden olabilir [2, 3, 4]. Bu nedenle, rulmanların sağlık izlemesi ve arıza teşhisi, endüstriyel ekipmanların güvenli ve kararlı çalışması için büyük önem taşımaktadır [5, 6, 7, 8]. Erken ve doğru arıza teşhisi, tahmine dayalı bakım stratejilerinde güvenilirliği, maliyet verimliliğini ve operasyonel sürekliliği sağlamak açısından kritik öneme sahiptir [9].

Geleneksel arıza teşhis yöntemleri, büyük ölçüde insan uzmanlığına ve zaman alıcı manuel öznitelik çıkarma süreçlerine

dayanmaktadır. Bunlar arasında istatistiksel göstergeler (örneğin, kurtosis, çarpıklık) ile hızlı Fourier dönüşümü (FFT), zarf analizi ve dalgacık dönüşümleri gibi klasik sinyal işleme teknikleri bulunmaktadır [10, 11]. Etkili olmakla birlikte, bu yaklaşımlar doğaları gereği alan bilgisine bağımlılıkları nedeniyle sınırlıdır ve farklı çalışma koşullarına veya makine tiplerine genelleme yapamayabilir. Yapay zeka (YZ) ve makine öğrenmesindeki (ML) [12], özellikle de derin öğrenmedeki [13] son gelişmeler, durum izleme için güçlü araçlar sunarak ekipman arızalarının gerçekleşmeden önce tahmin edilmesine olanak sağlamıştır. Çoğu uygulamada, ML yöntemleri büyük miktarda geçmiş veriden girdileri sınıflandırmayı öğrenir. Bununla birlikte, makine durum izleme alanında, izlemeye değer çoğu makinede arızanın yüksek maliyeti nedeniyle arıza verileri genellikle kıttır [14]. Bu kıtlık, yapay zeka sisteminin eğitim veri kümesinde doğru performans gösterip farklı çalışma koşulları altında elde edilen veriler, bir filodaki farklı makineler arasındaki veriler veya laboratuvardan sahaya taşındıktan sonraki veriler gibi diğer verilere genelleme yapamaması şeklinde önyargılı bir davranış sergilemesine yol açabilir. Tipik ML yöntemlerinin bir diğer dezavantajı, karar verme süreçleri hakkında hiçbir bilgi sunmayan kara kutu doğalarıdır ve bu da yorumlanabilirlik ve açıklanabilirlik eksikliğine neden olur [15, 16]. Bu sınırlama, özellikle makine durum izlemede yaygın olan güvenlik kritik alanlarda uygulamalarını kısıtlamaktadır.

Buna karşılık, derin öğrenme mimarileri, özellikle tek boyutlu evrişimsel sinir ağları (1D CNN'ler), ham titreşim sinyallerinden doğrudan hiyerarşik öznitelikleri otomatik olarak öğrenerek el yapımı öznitelik mühendisliğine olan ihtiyacı ortadan kaldırdıkları için öne çıkmaktadır [1, 9, 17]. Zaman serisi verileri üzerinde uçtan uca öğrenme gerçekleştirerek, 1D CNN'ler ön işleme sırasında değerli bilgilerin kaybedilmesi riski olmadan hem yerel hem de küresel zamansal bağımlılıkları yakalar, bu da onları gerçek zamanlı endüstriyel uygulamalar için son derece uygun kılar [9]. Bununla birlikte, üstün performanslarına rağmen, bu modellerin opaklığı güvenlik kritik ortamlarda kullanımının önünde kritik bir engel olarak kalmaktadır.

Modelle ilgili zorlukların ötesinde, literatür, önerilen birçok yöntemin genellenebilirliğini ve gerçek dünyada uygulanabilirliğini tehdit eden çeşitli kritik sınırlamalar ortaya koymaktadır. Önde gelen sorunlardan biri, Case Western Reserve University (CWRU) veri kümesi gibi yaygın olarak kullanılan referans veri kümeleriyle ilgilidir. CWRU veri kümesi fiili bir standart haline gelmiş olsa da, deneysel kurulumu genellikle göz ardı edilen kısıtlamalar getirmektedir [17, 18, 19]. Örneğin, farklı motor yük koşulları altında kayıtlar sağlanmış olsa da, herhangi bir mekanik aktarım (örneğin, dişliler veya kayışlar) bulunmadığından, bu farklılıklar rulmanlardaki radyal yükteki gerçek değişikliklere karşılık gelmemektedir. Asıl etki, arıza tespit edilebilirliğine ihmal edilebilir bir etkisi olan, şaft hızında küçük bir azalmadır (yaklaşık %4). Ayrıca, bazı yayınlar radyal yük konumunu yanlış bildirmiş ve bu da fiziksel gerekçesi olmayan teşhis varsayımlarına yol açmıştır [20]. Diğer bir kritik sorun, eğitim ve test kümelerinin rastgele bölünmesi gibi hatalı bölümlemenin, aynı kaydın komşu pencerelerinin her iki kümeye birden düşmesine ve gerçek dünya genellemesini yansıtmayan, modellerin neredeyse mükemmel performans elde ettiği yapay olarak şişirilmiş doğruluğa yol açabileceği veri sızıntısıdır [21]. Bunu ele almak, güvenilir bir değerlendirme sağlamak için sızıntısız çapraz doğrulama stratejileri gerektirir. Ayrıca, birçok çalışma rulman arızalarını yalnızca ayrık arıza boyutlarına göre sınıflandırmaktadır (örneğin, CWRU veri kümesinde 0,007, 0,014 ve 0,021 inç) [20], ancak arıza tespit edilebilirliği yalnızca boyuta değil, aynı zamanda konuma (iç bilezik, dış bilezik, yuvarlanma elemanı), yüke, sistem dinamiklerine ve ortaya çıkan titreşim imzalarına da bağlıdır. Sürekli bir bozulma sürecini ayrık kategorilere zorlamak, arıza ilerlemesinin kademeli ve düzensiz olduğu gerçek dünya verilerine genelleme yapma yeteneğini azaltabilir [17, 20].

Bugüne kadar, ham zaman alanı titreşim sinyalleri üzerinde eğitilmiş 1D CNN modellerinin performansını, aynı verilerden elde edilen frekans alanı temsilleri üzerinde eğitilmiş modellerle doğrudan karşılaştıran ve farklı çalışma koşullarına genelleme performansını sızıntısız bir protokol altında inceleyen sistematik bir çalışma yapılmamıştır. Bu boşluk önemlidir, çünkü girdi temsilinin seçimi hem teşhis doğruluğunu hem de farklı çalışma koşullarına genelleme yeteneğini etkileyebilir. Özellikle, frekans alanı genlik temsili faz bilgisini dışlamasına rağmen, rulman arızalarına ilişkin karakteristik frekans bileşenlerini daha belirgin hale getirebilir. Buna karşın, ham zaman alanı sinyallerindeki gürültü ve istenmeyen bileşenlerin performans üzerindeki etkisi de dikkate alınmalıdır. Bu nedenle, zaman alanındaki bilgilerin frekans alanına dönüşüm sırasında kaybedilmesinin teşhis performansını zorunlu olarak düşüreceği ve zaman alanındaki gürültünün performans farkında belirleyici bir rol oynayıp oynamadığı varsayımlarının deneysel olarak incelenmesi gerekmektedir.

Bu çalışmada, ham zaman serisi ve Ayrık Fourier Dönüşümü (DFT) tabanlı frekans alanı genlik temsilleri üzerinde eğitilen 1D CNN modellerinin rulman arıza teşhisindeki performansları karşılaştırılmaktadır. Modeller, yük seviyeleri 0–2 kullanılarak eğitilmiş ve tamamen görülmemiş yük seviyesi 3 verileri üzerinde test edilmiştir. Böylece, modellerin eğitim sırasında görülmemiş bir çalışma koşuluna genelleme yetenekleri değerlendirilmiştir. Her deneysel yapılandırma üç farklı rastgele tohum kullanılarak tekrarlanmış ve elde edilen sonuçlar karşılaştırılmıştır. Ayrıca, ham zaman alanı sinyallerine gürültü giderme işlemi uygulanarak, zaman alanı temsilindeki performansın gürültünün azaltılması ile iyileştirilip iyileştirilemeyeceği incelenmiştir. Sonuçlar, DFT tabanlı frekans alanı genlik temsilinin görülmemiş çalışma koşullarında ham zaman alanı temsilinden daha iyi genelleme performansı gösterdiğini ortaya koymaktadır. Gürültü giderme sonrasında zaman alanı temsilindeki performans farkının ortadan kalkmaması, bu farkın yalnızca giderilebilir gürültüden kaynaklanmadığını desteklemektedir. Bu bulgu, faz bilgisinin dışlanmasının tek başına teşhis performansında bir kayba yol açmadığını ve frekans alanı genlik temsilinin rulman arıza teşhisinde etkili bir girdi temsili olarak kullanılabileceğini göstermektedir.

## 2. Veri Kümesi Tanımı

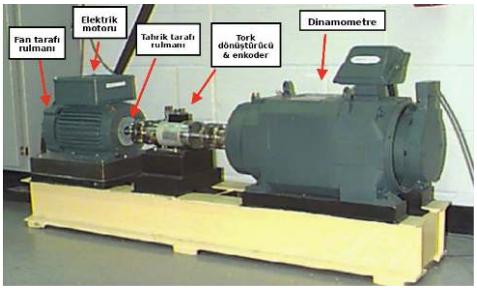
### 2.1. Rulman Arıza Özellikleri

Rulmanlardaki kusurlar (örneğin, yataklar veya yuvarlanma elemanları üzerindeki) tekrarlayan darbeler üretir ve rulman geometrisi ile şaft hızı tarafından belirlenen BPFO, BPFI, BSF ve FTF gibi karakteristik arıza frekanslarını oluşturur. Ham zaman

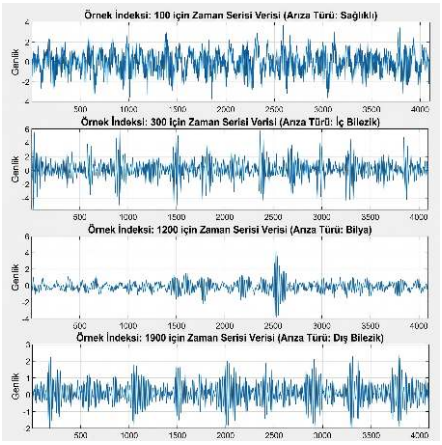
alanı sinyalinin frekans alanına dönüştürülmesi, bu periyodik imzaların daha belirgin hale getirilmesini sağlayabilir; ancak frekans alanı genlik temsili elde edilirken faz bilgisi dışlanmaktadır. Faz bilgisinin kaybının teşhis başarımını düşüreceği sıklıkla varsayılsa da [11], bu varsayımın geçerliliği bu çalışmada deneysel olarak incelenmektedir.

Zengin ve zorlu özellikleri nedeniyle Case Western Reserve University (CWRU) Bearing Data Center veri kümesi kullanılmıştır [19]. Veri kümesi, örnekleme frekansı ve arıza konumuna göre dört ana gruba ayrılmıştır: 48 kHz temel (normal çalışma), 12 kHz tahrik ucu arızası, 48 kHz tahrik ucu arızası ve 12 kHz fan ucu arızası. Her grup, bilye kusurları, iç bilezik (IR) arızaları ve dış bilezik (OR) arızaları dahil olmak üzere birden fazla arıza türü içermektedir. OR arızaları, yük bölgesine göre konumlarına göre daha da ayrılır: merkezi (saat 6 konumu), dik (saat 3 konumu) ve karşıt (saat 12 konumu). Arıza boyutları 0.007 ile 0.028 inç arasında değişmekte olup, motor yükleri 0 ile 3 hp arasında değişmektedir.

Şekil 1, CWRU rulman test düzeneğini ve rulman bileşenlerini göstermektedir. Bu çalışmada kullanılan tahrik ucu rulmanı bir SKF 6205-2RS JEM'dir; teknik özellikleri ve karakteristik arıza frekansları Tablo 1'te listelenmiştir. Sağlıklı, iç bilezik, bilye ve dış bilezik arızaları için temsili ham titreşim sinyalleri Şekil 2'te gösterilmiştir.




Şekil 1: CWRU rulman test düzeneği.




Şekil 2: Sağlıklı, iç bilezik, bilye ve dış bilezik arıza durumları için örnek ham titreşim sinyalleri.

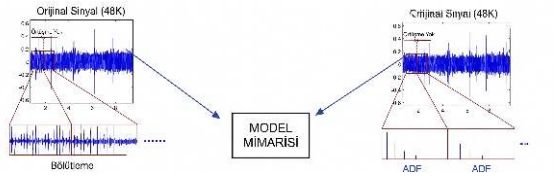
Tablo 1: Rulman detayları ve arıza frekansları [19].

| Özellik | Değer |
|---|---|
| **Düzenek üzerindeki konum** | Tahrik ucu |
| **Model numarası** | SKF 6205-2RS JEM |
| **BPFI** | 5.415 |
| **BPFO** | 3.585 |
| **FTF** | 0.3983 |
| **BSF** | 2.357 |

### 2.2. Veri Ön İşleme

Tahmine dayalı bakımda sonuçların kalitesi, ham sensör verisinin analiz için ne kadar etkili hazırlandığından önemli ölçüde etkilenir. Bu çalışmada, 48 kHz Tahrik Ucu Rulman Arızası alt kümesi kullanılmış ve aşağıdaki ön işleme adımları uygulanmıştır:

- **Bölümleme:** Sinyaller, 4096 örnekten oluşan sabit uzunlukta, üst üste binmeyen pencerelere bölünmüştür. İkinin kuvveti olan bu pencere boyutu, ayrık Fourier dönüşümünün verimli hesaplanmasını sağlamakta ve frekans alanı temsili için yeterli çözünürlük sunmaktadır.




Şekil 3: 4096 ardışık zaman noktasından oluşan tek bir örnek pencere.

- **Normalizasyon ve Kodlama:** Kararlı model yakınsaması için girdi özniteliklerini standartlaştırmak amacıyla Z-skoru normalizasyonu uygulanmıştır; veri sızıntısını önlemek için ölçeklendirme parametreleri yalnızca eğitim verisinden belirlenmiştir. Ayrıca, çok sınıflı sınıflandırma hedefiyle uyumluluğu sağlamak ve arıza türleri arasında yanlış sıralı ilişkiler çıkarılmasını önlemek için hedef arıza etiketleri bir-sıcak (one-hot) kodlama kullanılarak dönüştürülmüştür.

- **Filtreleme ve Yeniden Örnekleme:** Veri kümesi rulman teşhisi için yeterli frekans aralığına sahip yüksek doğrulukta kayıtlar sunduğundan, orijinal 48 kHz sinyaller ek filtreleme veya yeniden örnekleme yapılmadan kullanılmıştır.

### 2.3. Sınıflandırma Stratejisi ve Veri Bölümlenmesi

Arıza sınıflandırması, arıza boyutu ve çalışma koşullarındaki varyasyonlar bilinçli olarak birleştirilerek, kesin olarak sınıf etiketine (sağlıklı, bilye, iç bilezik veya dış bilezik) dayandırılmıştır. Bu yaklaşım, modeli, anlık çalışma durumundan bağımsız olarak temel mekanik arızayı tanımlamaya yönlendirir.

Görülmemiş çalışma koşullarına genelleme yeteneğini değerlendirmek için, motor yük seviyeleri 0, 1 ve 2 eğitim için kullanılmış, yük seviyesi 3 ise yalnızca test için ayrılmıştır. Belirsiz veya teşhis edilemeyen veri kümeleri, bir sonraki alt bölümde detaylandırıldığı gibi hariç tutulmuştur.

### 2.4. Teşhis Edilemeyen Veri Kümelerinin Hariç Tutulması

Birçok çalışma, CWRU içindeki belirli veri kümelerinin, mekanik montajın titreşimi gerçek arızadan daha fazla etkilediği belirsiz veya teşhis edilemeyen sinyaller içerdiğini vurgulamaktadır [18]. Özellikle, bazı arıza kayıtları yalnızca kısa geçici (transient) segmentler sergileyerek durağan olmayan (non-stationary) özellikler gösterirken, diğerleri arıza imzasını maskeleyen mekanik gevşeklikten muzdariptir. 48 kHz tahrik ucu alt kümesinin detaylı analizine dayanarak, Tablo 2'te listelenen kayıtlar eğitim aşamasından çıkarılmıştır. Yük 3 koşulunda bu listenin dışında bilye kaydı bulunmadığından, çapraz-koşul testinde bilye sınıfı söz konusu kayıtlarla temsil edilmiştir. Bu nedenle bilye sınıfına ait başarım değerleri, temsil edici kayıtların sınırlı teşhis edilebilirliği göz önünde bulundurularak değerlendirilmelidir.

Tablo 2: Teşhis edilemeyen veri kümeleri - Tahrik ucu rulman arızaları 48 kHz veri için filtrelenmiştir.

| Arıza Türü | Veri Kümeleri |
|---|---|
| **IR** | 174 |
| **Bilye** | 122, 123, 124, 125, 192, 228DE, 229DE |
| **OR merkezi** | 202FE, 204FE |
| **OR dik** | – |
| **OR karşıt** | – |

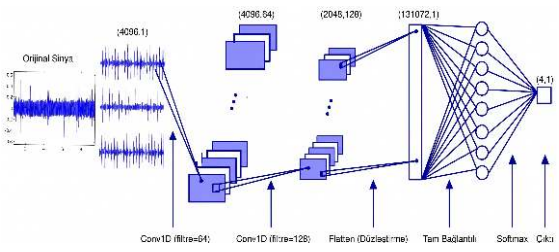
*Tahrik ucu rulman arızaları (48 kHz)

## 3. Model Mimarisi ve Eğitim Çerçevesi

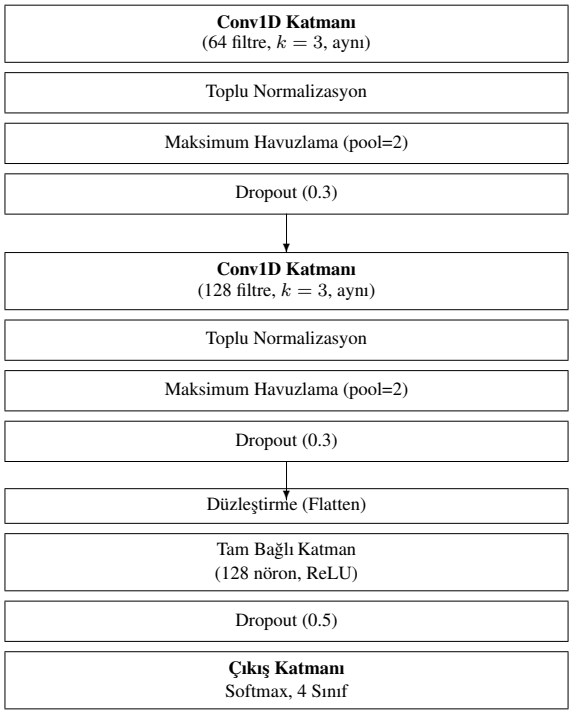
Rulman arıza teşhisi için derin öğrenme modelleri, yerel öznitelik çıkarımı için evrişimsel sinir ağları (CNN'ler), sıralı bağımlılıkları öğrenmek için uzun-kısa süreli bellek (LSTM) ağları ve uzun menzilli ilişkileri yakalamak için Transformer'lar gibi mimariler kullanmaktadır. Bu çalışmada, zaman serisi verilerini etkin bir şekilde işlemesi, manuel mühendislik gerektirmeden ilgili öznitelikleri otomatik olarak çıkarması ve titreşim tabanlı arıza teşhisi görevlerinde yaygın olarak kullanılması nedeniyle tek boyutlu evrişimsel sinir ağı (1D-CNN) benimsenmiştir.

Rulman durumlarını sınıflandırmak için uçtan uca bir 1D-CNN modeli kullanılmış ve böylece el yapımı öznitelik çıkarımına olan ihtiyaç ortadan kaldırılmıştır. Şekil 5'te gösterildiği gibi, modele girdi olarak iki farklı temsil kullanılmıştır: ham zaman alanı titreşim sinyali ve aynı sinyalden elde edilen DFT tabanlı frekans alanı genlik spektrumu. Her iki temsil de 4096 örnek uzunluğunda olup aynı 1D-CNN mimarisi ile işlenmiştir. Model, yerel ve zamansal özellikleri iki evrişim bloğu (Conv1D + toplu normalizasyon + maksimum havuzlama + dropout) üzerinden kendiliğinden öğrenir; ilk blok 3 çekirdek boyutuna sahip 64 filtre ve %30 dropout uygularken, ikinci blok benzer bir yapılandırmayla 128 filtre uygular. Çıkarılan öznitelikler düzleştirilir (flatten) ve 128 nörondan oluşan, ReLU aktivasyonlu tam bağlı bir katmandan geçirilir; ardından softmax çıkış katmanı sınıf olasılıklarını üretmeden önce %50 dropout uygulanır. Model, Adam optimize edicisi (öğrenme oranı= 0.001), 128 boyutunda yığın (batch) büyüklüğü ve en fazla 500 epoch kullanılarak eğitilmiştir. Kayıp fonksiyonu olarak kategorik çapraz entropi kullanılmış, aşırı uyumu (overfitting) azaltmak için dropout, toplu normalizasyon ve doğrulama izlemesi uygulanmıştır. Performans değerlendirmesi doğruluk, kesinlik (precision), duyarlılık (recall), F1-skoru ve karışıklık matrisi (confusion matrix) kullanılarak gerçekleştirilmiştir.

Hiperparametre seçimleri şu gerekçelere dayanmaktadır. Pencere boyutu 4096, ikinin kuvveti olduğundan FFT'yi verimli kılmakta ve yeterli frekans çözünürlüğü sağlamaktadır. İki evrişim bloğunda filtre sayısının 64'ten 128'e çıkarılması giderek daha soyut özniteliklerin öğrenilmesini hedefler; küçük çekirdek boyutu ($k$=3) titreşim sinyallerinde yerel desenleri yakalamak için yaygın bir seçimdir. Toplu normalizasyon eğitimi kararlı ve hızlı kılar; maksimum havuzlama (pool=2) boyutu azaltır ve küçük kaymalara dayanıklılık katar. Dropout değerleri (evrişimde %30, yoğun katmanda %50) özellikle büyük tam bağlı katmandaki aşırı uyumu sınırlamak için seçilmiştir. Adam optimize edici (öğrenme oranı= 0.001) ve 128'lik yığın boyutu, kararlılık ile hız arasında dengeli ve yaygın kabul görmüş varsayılan değerlerdir. Z-skoru normalizasyonu yalnızca eğitim verisinden hesaplanıp test kümesine uygulanarak veri sızıntısı önlenmiştir. Eğitim en fazla 500 epoch ile sınırlandırılmış; ancak testten bağımsız bir doğrulama kümesindeki başarım izlenerek en iyi genelleyen epoch'un ağırlıkları seçilmiştir. Son olarak, tek bir koşunun rastgele başlangıçtan kaynaklanan değişkenliğini hesaba katmak için her yapılandırma üç farklı tohumla tekrarlanmış ve sonuçlar ortalama±standart sapma olarak raporlanmıştır.




Şekil 4: Rulman Arıza Teşhisi için 1D-CNN Model Mimarisi.




Şekil 5: Önerilen 1D CNN Modelinin Mimarisi

## 4. Deneysel Sonuçlar

1D-CNN modelinin sınıflandırma başarımı; doğruluk, kesinlik (precision), duyarlılık (recall), F1-skoru ve karışıklık matrisleri ile değerlendirilmiştir. Tüm modeller yalnızca yük 0–2 kayıtlarıyla eğitilmiş ve tamamen görülmemiş yük 3 kayıtlarında (çapraz-koşul) test edilmiş, her yapılandırma üç farklı rastgele tohum ile tekrarlanmıştır. Tablo 3, ham zaman alanı, gürültü giderilmiş ham ve DFT genlik spektrumu temsilleri için ortalama±standart sapma değerlerini raporlamaktadır.

Tablo 3: Ham, gürültü giderilmiş ham ve DFT temsilleri için görülmemiş yük 3 test kümesindeki başarım (%; üç tohum: ortalama±std).

| Metrik (%) | Ham | Ham (denoising) | DFT |
|---|---|---|---|
| Doğruluk | 74.2±1.7 | 73.3±5.9 | **90.2±1.7** |
| Kesinlik (P) | 78.1±1.3 | 69.5±9.0 | **90.5±0.9** |
| Duyarlılık (R) | 73.6±2.3 | 72.7±5.8 | **92.9±1.6** |
| F1-skoru | 73.1±2.5 | 66.4±10.4 | **91.5±1.3** |

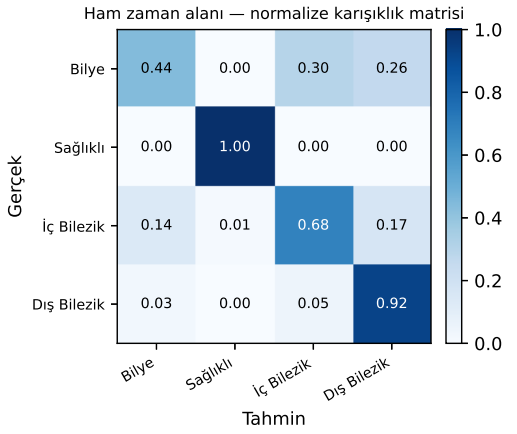
DFT genlik spektrumu üzerinde eğitilen model, görülmemiş yük 3 kayıtlarında %90.2 test doğruluğu ile ham zaman alanı modelini (%74.2) yaklaşık 16.0 yüzde puanı geride bırakmıştır; bu fark üç tohum boyunca standart sapmalarla tutarlıdır. Bu sonuç, "frekans alanına geçişte bilgi kaybı başarımı düşürür" yaygın varsayımının, en azından bu arıza-türü sınıflandırma görevi ve CWRU verisi için geçerli olmadığını göstermektedir.

Bu sonucun olası bir açıklaması şudur: DFT genlik spektrumu alınırken atılan faz bilgisi, darbelerin pencere içindeki kesin zamanlamasını taşır ve arıza türünün belirlenmesi için çoğunlukla gereksiz bir bileşendir. Arıza türünü ayırt eden bilgi (karakteristik arıza frekansları ve harmonikleri) ise genlik spektrumunda yer alır. Genlik spektrumunun tanım gereği zaman kaymasına duyarsız olması, bu temsilin görülmemiş çalışma koşullarına daha kararlı biçimde genellemesine katkıda bulunuyor olabilir. Bununla birlikte, bu mekanizmayı tek başına yalıtan ayrı bir deney yapılmadığından, söz konusu açıklama kesin bir neden değil, gözlemlenen sonucu destekleyen olası bir yorum olarak sunulmaktadır.

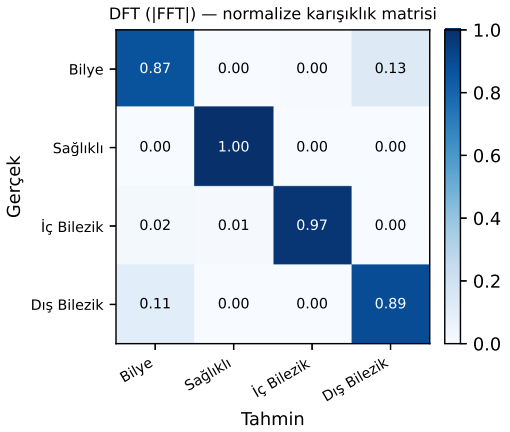
Bu açıklamayı sınamak için ham sinyallere dalgacık (db4, yumuşak eşik) tabanlı gürültü giderme uygulanarak aynı değerlendirme tekrarlanmıştır. Gürültü giderme ham modelin başarımını artırmamış (%73.3; ham ile aynı düzeyde ve daha değişken) ve DFT ile aradaki farkı kapatmamıştır. Bu da farkın giderilebilir gürültüden değil, temsilin yapısal özelliğinden (kaymaya duyarsızlık) kaynaklandığını desteklemektedir.

Sınıf bazında incelendiğinde (Şekil 6 ve 7), ham model özellikle bilye arızasını zayıf yakalamakta (duyarlılık %44) ve iç bilezik arızalarını kısmen karıştırmaktadır (%68); buna karşılık DFT modeli tüm arıza türlerinde daha dengeli ve yüksek duyarlılık elde etmektedir (bilye %87, iç bilezik %97, dış bilezik %89, sağlıklı %100).

Literatürde CWRU üzerinde rastgele bölünmeyle %99–100 civarı doğruluklar sıkça raporlanır; ancak bu değerler çoğunlukla aynı kaydın komşu pencerelerinin hem eğitime hem teste düşmesinden doğan veri sızıntısını içerir ve gerçek genellemeyi yansıtmaz [21, 20]. Bu çalışmada benimsenen sızıntısız, çapraz-yük protokolü daha gerçekçi bir başarım tahmini sunmaktadır.




Şekil 6: Ham zaman alanı modelinin görülmemiş yük 3 test kümesindeki normalize edilmiş karışıklık matrisi.




Şekil 7: DFT genlik spektrumu modelinin görülmemiş yük 3 test kümesindeki normalize edilmiş karışıklık matrisi.

## 5. Sonuç

Bu çalışma, rulman arıza teşhisinde ham zaman alanı ve DFT genlik spektrumu temsillerini, veri sızıntısını önleyen ve çapraz-çalışma koşulu genellemesini değerlendiren bir protokol altında, üç farklı rastgele tohum kullanarak sistematik olarak karşılaştırmıştır. Yaygın beklentinin aksine, görülmemiş çalışma koşullarında frekans alanı genlik temsili, ham zaman alanı temsiline kıyasla daha yüksek test başarımı sağlamıştır (%90.2'ye karşı %74.2). Ham sinyallere uygulanan gürültü giderme işlemi ise zaman alanı temsilindeki performans farkını kapatmamıştır. Bu bulgu, değerlendirilen çapraz-çalışma senaryosunda frekans alanı temsilinin farklı çalışma koşulları arasındaki değişimlere karşı daha kararlı bir temsil sağlayabildiğini göstermektedir.

Elde edilen sonuçlar, yalnızca girdi temsilinin değil, aynı zamanda değerlendirme protokolünün de derin öğrenme tabanlı arıza teşhis sistemlerinin genelleme performansı açısından belirleyici olduğuna işaret etmektedir. Özellikle, eğitim ve test verilerinin farklı çalışma koşullarından seçildiği çapraz-koşul değerlendirmesi, modelin koşul değişimine karşı dayanıklılığını daha doğrudan incelemeye olanak sağlamaktadır. Bununla birlikte, yüksek sınıflandırma başarımı tek başına modelin fiziksel olarak anlamlı kararlar verdiğini garanti etmemektedir. Bu nedenle, frekans alanı temsillerinin sağladığı avantajın yanı sıra, bu temsiller üzerinden elde edilen model açıklamalarının güvenilirliği ve fiziksel anlamlılığı da ayrıca ele alınmalıdır. Ayrıca, DFT genlik temsili ham zaman alanı temsilinden yalnızca faz bilgisinin dışlanmasıyla değil, aynı zamanda evrişimin uygulandığı eksenin değişmesiyle de ayrılmaktadır. Bu nedenle gözlenen performans farkının ne kadarının faz bilgisinin dışlanmasından, ne kadarının temsil ekseninin değişmesinden kaynaklandığı mevcut deney düzeneğiyle ayrıştırılamamaktadır; bu ayrımın eşleştirilmiş bir ablasyon düzeneğiyle incelenmesi devam eden çalışmamızın kapsamındadır.

Gelecek çalışmalar, bu bildiride elde edilen bulguların daha ileri taşınmasını hedeflemektedir. Devam eden ve bir dergi çalışması olarak yürütülen araştırmamızda, frekans alanında açıklanabilirliğin sağlanmasına odaklanılmaktadır. Bu kapsamda, 1D-CNN mimarisine gömülü DFT tabanlı bir yorumlanabilirlik katmanı ile SHAP analizi birleştirilerek model kararlarının karakteristik rulman arıza frekanslarıyla (BPFO, BPFI, BSF) ilişkisi incelenmekte ve üretilen açıklamaların fiziksel tutarlılığı değerlendirilmektedir. Ayrıca, önerilen sızıntısız çapraz-koşul değerlendirme protokolünün farklı çalışma koşulları ve ek kıyaslama veri kümeleri üzerinde genişletilerek yöntemin genelleme başarımının daha kapsamlı biçimde sınanması ve elde edilen sonuçların literatürdeki güncel yaklaşımlarla karşılaştırılması hedeflenmektedir. Bu doğrultuda, frekans alanı temsilleri ve açıklanabilir yapay zekâ yöntemlerinin rulman arıza teşhisindeki güvenilirliğini ve fiziksel anlamlılığını daha kapsamlı biçimde değerlendirmek amaçlanmaktadır.

## 6. Kaynakça

[1] Eren, L., Ince, T., & Kiranyaz, S. (2019). A generic intelligent bearing fault diagnosis system using compact 1D CNN classifier. *IEEE Transactions on Industrial Informatics, 15*(5), 2841–2850.

[2] Wang, B., Lei, Y., Li, N., & Li, N. (2020). A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings. *IEEE Transactions on Reliability, 69*(1), 401–412.

[3] Kong, X., Cai, B., Yu, Y., Yang, J., Wang, B., Liu, Z., Shao, X., & Yang, C. (2025). Intelligent Diagnosis Method for Early Faults of Electric-Hydraulic Control System Based on Residual Analysis. *Reliability Engineering & System Safety, 261*, 111142.

[4] Shao, X., Cai, B., Zou, Z., Shao, H., Yang, C., & Liu, Y. (2025). Artificial Intelligence Enhanced Fault Prediction with Industrial Incomplete Information. *Mechanical Systems and Signal Processing, 224*, 112063.

[5] Feng, J., Yao, Y., Lu, S. X., & Liu, Y. (2021). Domain Knowledge-Based Deep-Broad Learning Framework for Fault Diagnosis. *IEEE Transactions on Industrial Electronics, 68*(4), 3454–3464.

[6] Xing, S. B., Lei, Y. G., Wang, S. H., Lu, N., & Li, N. P. (2022). A label description space embedded model for zero-shot intelligent diagnosis of mechanical compound faults. *Mechanical Systems and Signal Processing, 162*, 108082.

[7] Shao, H. D., Xia, M., Han, G. J., Zhang, Y., & Wan, J. F. (2021). Intelligent Fault Diagnosis of Rotor-Bearing System Under Varying Working Conditions With Modified Transfer Convolutional Neural Network and Thermal Images. *IEEE Transactions on Industrial Informatics, 17*(5), 3488–3496.

[8] Deng, M. Q., Deng, A. D., Zhu, J., Shi, Y. W., & Liu, Y. (2021). Intelligent fault diagnosis of rotating components in the absence of fault data: A transfer-based approach. *Measurement, 173*, 108548.

[9] Hong, S., & Kim, Y. (2023). Real-time bearing fault diagnosis using 1D convolutional neural networks with raw

vibration data. *Mechanical Systems and Signal Processing, 186*, 109849.

[10] Zhang, W., Li, C., Peng, G., Chen, Y., & Zhang, Z. (2018). A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load. *Mechanical Systems and Signal Processing, 100*, 439–453.

[11] Randall, R. B., & Antoni, J. (2011). Rolling element bearing diagnostics–A tutorial. *Mechanical Systems and Signal Processing*, 25(2), 485–520.

[12] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, A. K. Nandi, Applications of machine learning to machine fault diagnosis: A review and roadmap, Mechanical Systems and Signal Processing 138 (2020) 106587.

[13] F. Jia, Y. Lei, J. Lin, X. Zhou, N. Lu, Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data, Mechanical systems and signal processing 72 (2016) 303–315.

[14] Zhao, Z., Li, T., Wu, J., Sun, C., Wang, S., Yan, R., & Chen, X. (2020). Deep learning algorithms for rotating machinery intelligent diagnosis: An open source benchmark study. *ISA Transactions*, 107, 224–255.

[15] Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation*, 16, 45–74.

[16] Herwig, N., Borghesani, P., Smith, W., & Peng, Z. (2025). Signal processing- and physics-informed neural network for explainable bearing condition monitoring. *Mechanical Systems and Signal Processing*, 235, 112925.

[17] Neupane, D., & Seok, J. (2021). Bearing fault detection and diagnosis using deep learning: A comprehensive review. *IEEE Access, 9*, 122341–122364.

[18] Smith, W. A., & Randall, R. B. (2015). Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mechanical Systems and Signal Processing*.

[19] Case Western Reserve University Bearing Data Center website, 2021. `https://engineering.case.edu/bearingdatacenter/welcome`.

[20] Hendriks, J., Dumond, P., & Knox, D. A. (2022). Towards better benchmarking using the CWRU bearing fault dataset. *Mechanical Systems and Signal Processing*.

[21] Rosa, J. L., Garcia, A. M., & Sanz, R. (2022). Data leakage in the CWRU bearing dataset: Consequences and solutions. *Measurement, 190*, 110695.