Sensör Füzyonu ve Derin Öğrenme Teknikleri Kullanarak Hava Araçlarında Anomali Tespiti

Yıldırım ÇAT[1,2*], Prof. Dr. Kemal FİDANBOYLU[1**]

[1]*Bursa Uludağ Üniversitesi, Bilgisayar Mühendisliği Bölümü, Bursa, TÜRKİYE*

[2]*Turkish Aerospace Industry, TAI-Uludağ University R&D Center, Bursa, TÜRKİYE*

Özet

Havacılık endüstrisinde kullanılan bileşenler, belirli bir süre kullanıldıktan sonra performas kaybı yaşamakta ve bozulmaktadır. Bu bileşenler önceden planlanmış periyodik bakımlara tabi tutularak tekrar son kullanıcının kullanımına sunulur. Ancak, bileşenlerin çok sık bakıma alınması ve sürekli olarak değiştirilmesi, sistem performansında ve hizmet kalitesinde istenmeyen düşüşlere ve bekleme sürelerinin artmasına neden olmaktadır. Bu dezavantajları gidermek için Yapısal Sağlık İzleme (SHM) ve Kestirimci Bakım (PdM) teknikleri yaygın olarak kullanılmaktadır. Bu tekniklerde yaygın olarak kullanılan yaklaşım anomali tespitidir. Anomali tespitinde, bir sistemdeki genel kurala uymayan ve belirli bir konu hakkında bilgi sağlayan istisnai örneklerin bulunması hedeflenmektedir. Bileşenlerde zamanla meydana gelen ve istenmeyen arızalar, anomali tespiti yaklaşımları kullanılarak teşhis edilebilmektedir. Anomali tespitinde kullanılan yöntemlerden biri sensör füzyonudur. Sensör füzyonu, birden fazla sensörden gelen verilerin birleştirilmesi ve analiz edilmesi işlemidir. Havacılık sektöründe, sensör füzyonunun kullanılmasıyla hata payı daha düşük verilerin toplanarak yeni bir veri seti üretilebilir. Elde edilen füzyonlanmış veriler üzerinde yapılan anomali tespiti süreci de daha doğru sonuçlar üretmektedir. Bu süreçte, Kalman Filtresi veya Demspter-Shafer Teorisi yaygın olarak kullanılan sensör füzyonu teknikleri olarak öne çıkmaktadır.

Bu bildiride, NASA tarafından paylaşılan bir çalışmada, gürültüyü ortadan kaldırmak ve gürültülü sinyallerin zayıf özniteliklerini çıkarmak için rulmanlardan oluşturulan bir test düzeneğinden elde edilen zaman serisi verileri üzerinde anomali tespiti yapılmıştır. Geliştirilen sistemde rulmanlardan elde edilen ham titreşim verileri ön işleme tabi tutulduktan sonra her bir veri noktası kalman filtresine girdi olarak verilmiş ve yeni çıktı değerleri tahmin edilmiştir. Tahmin edilen veriler Autoencoder modelini beslemek için kullanılmıştır. Model eğitimi sürecinde optimizasyon algoritması olarak Gradyan İniş algoritmasının gelişmiş bir varyantı olan Adam algoritması kullanılmıştır. Eğitilen model kullanılarak eşik değer aşma tabanlı anomali tespiti yapılmıştır.

Anahtar Kelimeler: *Anomali Algılama, Yapısal Sağlık İzleme, Tahmini Bakım, Uçak, Havacılık Endüstrisi, Derin Öğrenme, Sensör Füzyonu, Kalman Filtresi, Demspter-Shafer Teorisi, Gradyan İniş Algoritması, Adam Algoritması*

Anomaly Detection in Aircraft Using Sensor Fusion and Deep Learning Techniques

Abstract

Components used in the aviation industry lose performance and deteriorate after a certain period of use. These components are subjected to pre-planned periodic maintenance and are made available to the end user again. However, frequent maintenance and constant replacement of components cause undesirable decreases in system performance and service quality and increased waiting times. To overcome these drawbacks, Structural Health Monitoring (SHM) and Predictive Maintenance (PdM) techniques are used. The commonly

*Contact email: 502331024@ogr.uludag.edu.tr

used approach in these techniques is anomaly detection. In anomaly detection, the aim is to find exceptional examples that do not comply with the general rule in a system and provide information about a specific subject. Faults that occur in components over time can be diagnosed using anomaly detection approaches. A common method used in anomaly detection is sensor fusion. Sensor fusion is the process of combining and analyzing data from multiple sensors. In the aviation industry, by using sensor fusion, a new data set can be produced by collecting data with a lower margin of error. The anomaly detection performed on the fused data obtained produces more accurate results. In this process, Kalman Filter or Dempster-Shafer Theory stand out as widely used sensor fusion techniques.

In this paper, in a study shared by NASA, anomaly detection was performed on time series data obtained from a test rig created from bearings to eliminate noise and extract weak features of noisy signals. In the developed system, after the raw vibration data obtained from the bearings were pre-processed, each data point was given as input to the Kalman filter and new output values were estimated. The predicted data was used to feed the Autoencoder model. During the model training process, the Adam algorithm, an advanced variant of the Gradient Descent algorithm, was used as the optimization algorithm. Threshold exceedance-based anomaly detection was performed using the trained model.

Keywords: *Anomaly Detection, Structural Health Monitoring, Predictive Maintenance, Aircraft, Aviation Industry, Deep Learning, Sensor Fusion, Kalman Filter, Demspter-Shafer Theory, Gradient Descent Algorithm, Adam Algorithm*

1 Giriş

Endüstride yararlanılan bileşenler, kullanıma bağlı olarak zamanla bozulmakta ve genellikle sabit zaman aralıklarında planlı bakıma tabi tutulmaktadırlar. Bu yaklaşımın dezavantajları; bileşenlerin çok sık değiştirilmesi, sağlıklı parçanın değiştirilmesi veya bir bileşenin arıza sonrası değiştirilmesidir. Bu dezavantajlar, verimliliğin ve ürün kalitesinin düşmesine veya sistemde istenmeyen bekleme sürelerine yol açmaktadır [1]. Sık gerçekleşen bakımlar veya doğru planlanmayan bakım süreçleri maliyet artışına sebep olmaktadır [2]. Sahip olduğu dezavantajlar nedeniyle bu yaklaşım yerini Yapısal Sağlık İzleme (Structural Health Monitoring, SHM) ve Kestirimci Bakım (Predictive Maintenance, PdM) uygulamalarına bırakmaktadır. SHM ve PdM uygulamalarında proaktif önlemlerin alınmasında kullanılan bir yaklaşım ise anomali tespitidir. Sistemdeki herhangi bir bileşende oluşabilecek anomali veya hatanın erkenden teşhis edilmesi, sistemin bakım planında önemli seviyede değişikliklere gidilmesine yol açabilir. Bu gibi bir senaryoda etkin bir anomali tespiti yönteminin kullanılması sistemin sürdürülebilirliği ve bakım planı için önem arz etmektedir.

Bileşenlerde meydana gelen anomaliler bileşenlerin zamanla bozulmalarına yol açar. Bu bozulmaların erken evrede belirlenmesi anomali tespiti yöntemleri ile gerçekleştirilmektedir. Anomali tespiti ile bir sistemdeki genel kurala uymayan ve belirli bir konu hakkında bilgi sağlayan istisnai örneklerin bulunması hedeflenmektedir [3]. Anomaliler üç başlıkta gruplandırılabilir; bir sistemde sadece tek başına bulunan Nokta Anomali, belirli bir bağlamda anomali olup diğer bağlamlarda anomali olarak sayılmayan, Bağlamsal Anomali ve grup halinde anomali olarak bulunan Toplu Anomali [4]. Literatürde bu anomali türlerinin tespiti için sensör füzyonu ile Makine Öğrenmesi (Machine Learning, ML) ve Derin Öğrenme (Deep Learning, DL) teknikleri kullanılmıştır.

Sensör füzyonu, ML ve DL yöntemlerinin detayları ve literatürdeki uygulamaları ikinci bölümde paylaşılacaktır. Aynı bölümde, bu çalışmada kullanılan ve havacılık alanında yapılan birçok çalışmada kullanılan Cincinnati Üniversitesi Akıllı Bakım Sistemleri Merkezi (University of Cincinnati Intelligent maintenance Systems, IMS) veri setinin ve tekniklerin açıklaması yapılacaktır. Üçüncü bölümde geliştirilen hibrit model sunulacaktır. Dördüncü bölümde yürütülen deneylerin sonuçları tartışıldıktan sonra son bölümde ise yapılan çalışmanın sonuçları sunulacaktır.

2 Literatür araştırması

2.1 Hava araçlarında anomali tespitinde kullanılan füzyon ve derin öğrenme teknikleri

Havacılık endüstrisinin hızla gelişmesine paralel olarak bu alanda kullanılan mekanik ve yazılımsal tüm bileşenler de daha kompleks hale gelmiştir. Bu

durumun bir sonucu olarak, bu alan üzerinde yapılan tüm araştırmalar ve deneyler daha girift hale gelmiştir. Bu çalışmanın da konusu olan ve havacılık endüstrisinde oldukça yaygın bir şekilde uygulanan anomali tespiti yöntemleri, özellikle sensör füzyonu ve DL tekniklerinin entegre edilmesiyle büyük oranda gelişim göstermiştir. Başlıca kullanılan sensör füzyonu teknikleri Dempster-Shafer teorisi ve Kalman filtresi olarak sıralanabilirken; Uzun Kısa Süreli Bellek (Long Short-Term Memory, LSTM), Otokodlayıcı (Autoencoder, AE) ve Evrişimsel Sinir Ağları (Convolutional Neural Networks, CNN) DL modelleri de anomali tespiti uygulamalarında ilk sıralarda tercih edilmektedir.

2.1.1 Dempster-Shafer teorisi (DST)

Dempster-Shafer teorisi (DST), Bayes teorisinin bir genellemesidir ve Bayes teorisinin eksikliklerini gidermek için geliştirilmiştir [5]. Teori, kısaca belirsizliği yönetmek, ML ve DL'de çoklu kaynaktan elde edilen bulguları kombine etmek için kullanılmaktadır. Her bir olasılık kümesine ağırlık atayarak belirsizliğin daha açıklayıcı bir temsilinin oluşmasını ve birbiriyle çatışan bulguların üstesinden gelinmesini sağlayarak belirsizliğin hakim olduğu ortamlarda entegre edildiği ML veya DL modelinin daha tutarlı ve daha güvenilir olmasını sağlamaktadır.

Garcia ve ark. [6], havacılık sistemlerinde meydana gelen karmaşık yapıdaki anomalilerin tespitinde kullanılan ve insan tarafında yorumlanabilirliği zayıflatma pahasına geliştirilen DL modellerine alternatif bir yaklaşım önermişlerdir. Yeni yaklaşımda sensör füzyonu icra edilerek karmaşık sınıflandırıcılara olan bağımlılığın azaltılması amaçlanmıştır. Geleneksel eşik-aşımı yöntemleriyle elde edilen yorumlanabilirlik seviyesi korunurken sıralı örüntüler ve bunların dağılımları kullanılarak yeni bir tahmin modeli üzerinde çalışılmıştır. Önerilen yaklaşım, topluluk öğrenmesi ile eğitilen sınıflandırıcıların tahminlerindeki belirsizliği minimize etmek için DST'yi kullanmaktadır. Birden fazla değişkenin kombinasyonu, belirli anomalilerin tespiti için dikkate alınmıştır ve elde edilen sonuçlar, DST tabanlı bu yeni yöntemin çoğunluğun gücü yaklaşımına göre daha iyi performansa sahip olduğu görülmüştür.

2.1.2 Uzun kısa süreli bellek (LSTM)

LSTM modeli, Yinelemeli Sinir Ağı'nın (Recurrent Neural Network, RNN) bir türüdür ve geleneksel ileri-beslemeli sinir ağlarından farklı olarak önceki zaman adımlarındaki bilgileri de hatırlayabilmektedir. Standart bir RNN ağı, çıktıyı üretirken mevcut adımdaki girdiyle beraber bir önceki adımda elde edilen çıktıyı da kullanır. Böylece, mevcut çıktı hesaplanırken önceki adımdaki bilgiler de göz önünde bulundurulur [7]. Ancak, RNN ağının bir eksikliği sadece yakın zamandaki geçmiş bilgileri tekrar kullanabilmesidir. Uzak geçmişteki bilgiler tamamen unutulmakta ve mevcut çıktının hesabında kullanılamamaktadır, yani diğer bir deyişle Kaybolan Gradyan Sorunuyla (Vanishing Gradient Problem) karşılaşılmaktadır [8]. LSTM, RNN'deki bu eksikliği gidermek amacıyla geliştirilmiştir ve hem kısa hem de uzun süreli geçmişteki bilgiler mevcut çıktının üretimi sürecinde kullanılır. Mimarisinde girdi katmanı, gizli katman ve çıktı katmanını bulunduran model, RNN'den ayrı olarak gizli katmanı içerisinde özel bir LSTM hücresi (LSTM Cell) içermektedir [9]. Üretilen çıktıların unutulma oranı, hangi yeni bilgilerin ekleneceği veya güncelleneceği ve saklanan bilgilerin hangi oranda kullanılacağı bu özel hücre ile belirlenir. Bu haliyle, kompleks veri setleriyle çalışıldığı durumlarda, standart RNN modelinden çok daha iyi tahminleme kabiliyetine sahiptir. Özellikle zaman serisi verilerinde, sıralı zaman noktalarının birbirleriyle ilişkileri, yapılacak tahmin için kritik seviyede öneme sahip olduğundan LSTM modeli yaygın olarak kullanılmaktadır [10]. Zaman serisi veri setleri de havacılık endüstrisinde anomali tespiti, Durum İzleme (Condition Monitoring) ve PdM uygulamalarında üretilmektedir.

2018 yılında NASA tarafından Mars yüzeyine gönderilen Curiosity isimli uzay araştırma aracından ve Soil Moisture Active Passive (SMAP) uydusundan elde edilen telemetri verilerini kullanarak bir gözetimsiz anomali tespiti çalışması yürütülmüştür [11]. Çalışmada kullanılan her iki veri setinde de uzmanlar tarafından etiketlenmiş anomali verileri yer almaktadır. Yüksek karmaşıklığa sahip zamansal veya sıralı veriler karşısında iyi sonuç vermesi ve geçmiş verilerle güncel veriler arasındaki ilişkiyi öğrenebilmesi nedeniyle çalışmada, LSTM ağları kullanılmıştır. Her bir telemetri kanalı için tekli modeller oluşturulmuş ve ilgili kanaldaki değerlerin tahmini yapılmıştır.

2.1.3 Evrişimsel sinir ağları (CNN)

Evrişimsel Sinir Ağları (Convolutional Neural Networks, CNNs) bir çeşit yapay sinir ağı olmakla

birlikte insandaki görme merkezinden esinlenilerek geliştirilmiştir [12]. Görüntü işleme alanında yaygın olarak tercih edilen bu model ile keskinleştirme, kenar tespiti ve bulanıklaştırma gibi birçok uygulama icra edilmekte ve yüksek performans elde edilmektedir. Model; Evrişim (Convolution), Havuzlama (Pooling) ve Tam Bağlantılı (Fully-Connected) olmak üzere üç katmandan oluşmaktadır. Girdi, Evrişim katmanında Denklem (1) ile filtreleme sürecine tabi tutulup sonraki katman olan Havuzlamaya gönderilir [13]. Bu katmanda sadece önemli özniteliklerin elde edilmesi ve maliyetin düşürülmesi amacıyla bir boyutsal indirgeme yapılmaktadır. Bu işlem sayesinde hesaba katılacak sinir ağı sayısı ve parametre sayısı azaldığı için aşırı öğrenme (overfitting) problemiyle karşılaşma olasılığı da düşürülmüş olur [14]. Tam Bağlantılı katmanda ise girdi Çok Katmanlı Algılayıcı (Multilayer Perceptron, MLP) ağına gönderilmek üzere sütun vektörü olarak düzleştirilir.

$$\begin{bmatrix} x_{1,1} & x_{1,2} \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} \cdots & x_{m,n} \end{bmatrix} * \begin{bmatrix} y_{1,1} & y_{1,2} \cdots & y_{1,n} \\ \vdots & \ddots & \vdots \\ y_{m,1} & y_{m,2} \cdots & y_{m,n} \end{bmatrix} = \sum_{i=1}^{m} \sum_{j=1}^{n} x_{i,j} y_{i,j} \quad (1)$$

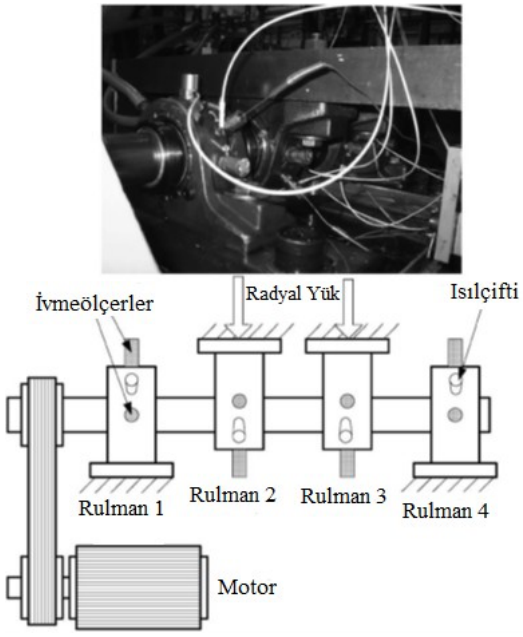
Wang ve ark. [15], rulmanlarda oluşabilecek anomalilerin önceden tespit edilebilmesi amacıyla ham titreşim sinyallerini görüntülere dönüştürmeyi önermişlerdir. Sinyal verilerini görüntüye dönüştürmede kullanılan bir yöntem olan Markov Transition Field (MTF) kullanılarak, titreşim sinyalleri görüntülere dönüştürüldükten sonra CNN ile özniteliklerin elde edilmesi ve anomali tespitinin yapılması amaçlanmıştır. Yöntemin performansını test etmek için Case Western Reserve University (CWRU) ve Mechanical Failure Prevention Technology (MFPT) veri setleri kullanılmıştır. İki veri seti ile yapılan deneylerin sonucunda yüksek doğruluk oranı elde edilmiştir.

2.2 Hava araçlarında anomali tespitinde kullanılan yöntemler

2.2.1 Veri Seti

Qiu ve ark. [16], sinyallerin gürültülerinin giderilmesi ve zayıf özniteliklerinin çıkarılması için yaptıkları çalışmada, geliştirdikleri Dalgacık Filtre (Wavelet Filter, WF) yönteminin performansını doğrulamak amacıyla rulmanlardan oluşan bir test düzeneği oluşturmuşlardır [16]. Oluşturulan deney düzeneğinde Şekil 1'de görüldüğü gibi bir mil üzerine dört adet çift sıralı rulman yerleştirilmiştir. Milin dönme hızı dakikada 2,000 devir olarak ayarlanmıştır. Mil üzerinde kurulan bir yay mekanizması ile toplamda yaklaşık 2,721.5 kilogramlık bir yük eklenmiştir. Rulmanların titreşim verileri 20 dakikalık aralıklarla kaydedilmiştir. Veriler 20 KHz frekansla örneklenmiş ve toplamda 20,480 veri noktası elde edilmiştir. Yapılan testlerle toplanan titreşim verileri kullanılarak üç farklı veri seti oluşturulmuştur. Bu çalışmada da kullanılan, ikinci testte toplanan ve toplamda 984 adet örnekten oluşan veri setinde rulman 1'de dış bilezik hatası gözlemlenmiştir.

Wang ve ark. [17], Liu ve ark. [18], Ho ve ark. [19] yürüttükleri çalışmalarda rulmanlarda meydana gelebilecek anomalilerin ve arızaların teşhisi için IMS veri setini kullanmışlardır.



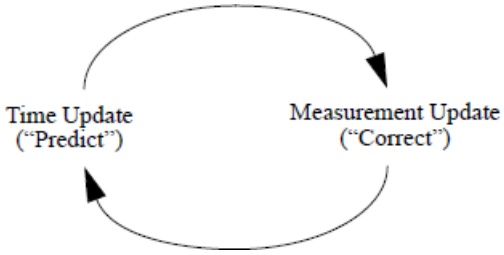
Şekil 1. Rulman test düzeneği [20]

2.2.2 Kalman filtresi

Kalman filtresi, dinamik sistemlerde zamanla değişen durumları izlemek ve tahmin etmek için geliştirmiştir. İlk olarak Rudolf E. Kálmán tarafından 1960 yılında tanıtılmıştır [21]. Kalman filtresi, sistemin durumunu modellemek için doğrusal bir dinamik sistem ve buna ek olarak gözlemler arasında gürültü bulunduğunu varsayar.

Temel olarak Şekil 2'de gösterildiği gibi iki aşamadan oluşur: Tahmin aşaması (Predict) ve Güncelleme (Correct) aşaması. Tahmin aşamasında, bir sonraki zaman adımı için bir önsel olasılık

(priori) tahmini elde etmek üzere mevcut durum ve hata kovaryans tahminleri zamanda ileriye doğru yöneltilir. Güncelleme adımlarında ise geri besleme işlemi yapılır.



Şekil 2. Kalman filtresi döngüsü [22]

Lee ve ark. [23], değişken kirlilik seviyelerinin varlığında bir hidrolik dişli pompasının kalan faydalı ömrünün (Remaining Useful Lide, RUL) tahminini içeren üç sensörlü bir füzyon tekniği önermiştir. Araştırmacılar, bu durumda üç farklı sensör veri akışından titreşim, akış hızı ve basınç sinyallerinden hata özelliklerini düzgün bir şekilde kaynaştırmak için Kalman filtresi tabanlı doğrusal bir model kullanmışlardır. Bu birleştirilmiş özellikler daha sonra Çift Yönlü LSTM (Bidirectional LSTM, BI-LSTM) ağı için bir girdi olarak kullanılmış ve güçlü bir RUL mimarisinin oluşturulmasıyla sonuçlanmıştır.

2.2.3 Gradyan iniş algoritması (GDA)

Gradyan İniş Algoritması (Gradient Descent Algorithm, GDA), türevlenebilir bir kayıp (maliyet) fonksiyonunun minimum değerini bulmak için geliştirilen yinelemeli bir yöntemdir [24]. Kayıp fonksiyonu genellikle parametreler (ağırlıklar) açısından türetilir ve bu türev (gradyan), fonksiyonun eğimini ve yönünü gösterir. GDA, parametrelerin değerlerini bu eğim boyunca adım adım ayarlayarak kayıp fonksiyonunu minimize etmeye çalışır. Denklem (2)'de algoritmanın standart matematik modeli verilmiştir. Burada $\theta_)$ ile mevcut parametre ifade edilirken; (θ), mevcut parametrenin kayıp fonksiyonunu temsil etmektedir. Gradyan alma işlemi nabla operatörü ∇ ile temsil edilmiş, genellikle 0-1 aralığında olan α ile öğrenme oranı temsil edilirken tüm bu işlemin sonucu da bir sonraki adımdaki parametreye yani, $\theta_{)*!}$ ifadesine eşitlenmiştir. Kayıp fonksiyonu minimize edilinceye kadar veya sistemde tanımlı belirli bir durdurma kriterine, maksimum iterasyon sayısına ulaşılması veya eşik değerinin aşılması, karşılaşılana kadar bu denklem işletilmeye devam edilir.

$$\theta_{t+1} = \theta_t - \alpha \nabla J(\theta) \qquad (2)$$

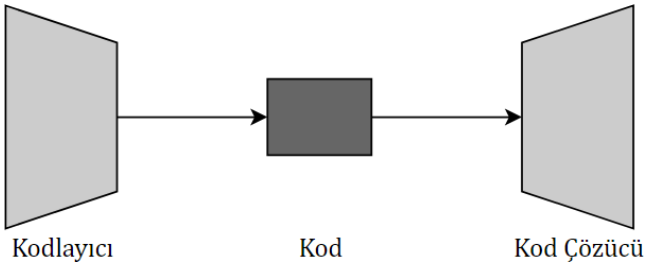
Öğrenme oranının manipüle edilmesiyle GDA'nın daha iyi sonuç veren varyantları geliştirilmiştir. DL uygulamalarında yaygın bir şekilde kullanılan Adam (Adaptive Moment Estimation) algoritması da bunlardan bir tanesidir. İlk olarak 2015 yılında geliştirilen bu algoritma [25], özellikle gradyan ve karesel gradyanın üstel bir hareketli ortalamasını hesaplar ve tıpkı α gibi iki farklı parametre ile bu hareketli ortalamaların bozunma oranlarını kontrol eder. DL uygulamalarında Adam algoritması diğer stokastik optimizasyon algoritmalarına göre daha yüksek performansa sahiptir ve daha iyi sonuçlar üretmektedir.

2.2.4 Otokodlayıcı (AE) modeli

AE, ilk defa Kramer tarafından tanıtılmış ve temelde bir yapay sinir ağıdır [26]. Geleneksel bir AE Şekil 3'te de görülebileceği üzere üç katmandan meydana gelmektedir: Kodlayıcı (Encoder), Kod veya Gizli Alan (Code, Latent Space) ve Kod Çözücü (Decoder). Kodlayıcı katmanında işletilen gizli fonksiyonla girdi üzerinde bir dönüşüm uygulanır ve sıkıştırılarak Gizli Alan katmanına aktarılır. Bu katmanda, sinir ağları kullanılarak girdiye olabildiğince benzer çıktı tahmin edilir. Sonrasında Kod Çözücü katmanının icra ettiği dönüşüm ile girdinin Yeniden Oluşturma (Reconstruction) süreci başlatılır. Bu süreç boyunca sıkıştırılan veriler eski boyutuna geri döndürülür. İdeal performans düzeyinde çalışan bir AE için girdi verileri ile çıktı verileri birbirine oldukça yüksek düzeyde benzemelidir. AE modelinin yaygın olarak kullanıldığı uygulamalardan birisi anomali tespitidir. AE'nin Kodlayıcı katmanına verilen girdi ile Kod Çözücü katmanının ürettiği çıktı arasındaki farka Yeniden Oluşturma Hatası (Reconstruction Error, RE) adı verilmektedir. Optimal çalışan bir AE'ye girdi olarak verilen normal (anomali olmayan) verinin girdi ve çıktı değerleri arasındaki RE değeri minimum seviyelerdeyken; girdi olarak verilen anomali verisinin girdi ve çıktı arasındaki RE değeri daha yüksek seviyelerde olmaktadır.

Bu anomali tespiti yönteminin matematiksel modelinde [27], Denklem (3)'te d fonksiyonu ile bahsedilen RE'nin derecesi, x ile girdi ve x' ile de yeniden oluşturulan x çıktısı temsil edilmektedir.

$$\mathrm{d}(x, x') \qquad (3)$$



Şekil 3. AE mimarisi

Denklem (4)'teki gibi bir kayıp fonksiyonu tanımlandığında: $\mathbb{E}_{x\sim\mu_{ref}}$ ifadesi, başlangıçta AE'nin yapması gereken süreci temsil etmek üzere girdi kümesi için bir referans olasılıksal dağılımı temsil ederken; $D_\theta(E_\emptyset(x)))$ ifadesi, Kod Çözücü'nün yaptığı yeniden oluşturma sürecini ifade eder. Burada, $E_\emptyset(x)$ her x girdisi için $\emptyset$ ve θ parametrelerine bağlı olarak Kodlayıcıyı temsil etmektedir.

$$L(\theta,\emptyset) := \mathbb{E}_{x\sim\mu_{ref}}[d(x, D_\theta(E_\emptyset(x)))] \quad (4)$$

İdeal AE için Denklem (5)'teki gibi bir optimizasyon süreci icra edilir. Burada, θ ve $\emptyset$ parametrelerine bağlı olan AE'nin kullandığı kayıp fonksiyonun çıktıyı en aza düşürmesi istenen durumdur. Yani, bu işlem gradyan iniş algoritması ile yapılmaktadır [28].

$$argmin_{\theta,\emptyset}\, L(\theta,\emptyset) \quad (5)$$

Singh ve ark. [29] dişli çarklarda bulunan rulmanların RUL değerlerini tahmin edebilmek için AE kullanarak gözetimsiz anomali tespiti çalışması yapmışlardır. NASA IMS veri setinde anomali etiketli verilerin olmamasından dolayı AE modeli seçilmiştir. Eğitimde oluşan maksimum hata eşik değeri olarak seçilmiştir. Deney sonuçlarına göre, sadece bir gizli katman kullanılarak oluşturulan AE modeli ile anomali tespitinde başarılı sonuçlar elde edilmiştir.

Ahmad ve ark. [30], döner mekanizmalardaki anomalilerin tespiti ve durumlarının izlenmesi için LSTM Tabanlı AE (LSTM-Based Autoencoder) modeli geliştirmişlerdir. Çalışmada NASA IMS veri seti ve 5 farklı rulman verisi içeren Endüstri (Industry) veri seti kullanılmıştır. Anomali tespit aşamasında, normal verilerden elde edilen RE değerlerinin anomali skorlarına dönüştürülmesiyle eşik değeri hesaplanmıştır. Yapılan deneyler sonucunda, her iki veri setindeki anomaliler de doğru bir şekilde tespit edilmiştir.

Li ve ark. [31], büyük endüstriyel mekanik ekipmanlardaki oluşabilecek anomalileri tespit etmek amacıyla AE ve Zamansal Evrişimsel Ağ (Temporal Convolutional Network, TCN)'ı kullanan yeni bir gözetimsiz yaklaşım önermişlerdir. AE sinyallerin RE değerini hesap etmek için kullanılırken; anomali verilerinin az olması problemini bertaraf edebilmek için TCN kullanılmış ve zaman serisi sinyalleri tahmin edilmiştir. Geliştirilen sistemin etkinliğini ölçebilmek için IMS veri seti üzerinde yapılan deneyler sonucunda Test 2 verileri için elde edilen metrikler geliştirilen yaklaşımın başarısını doğrulamıştır.

3 Uygulanan yöntem

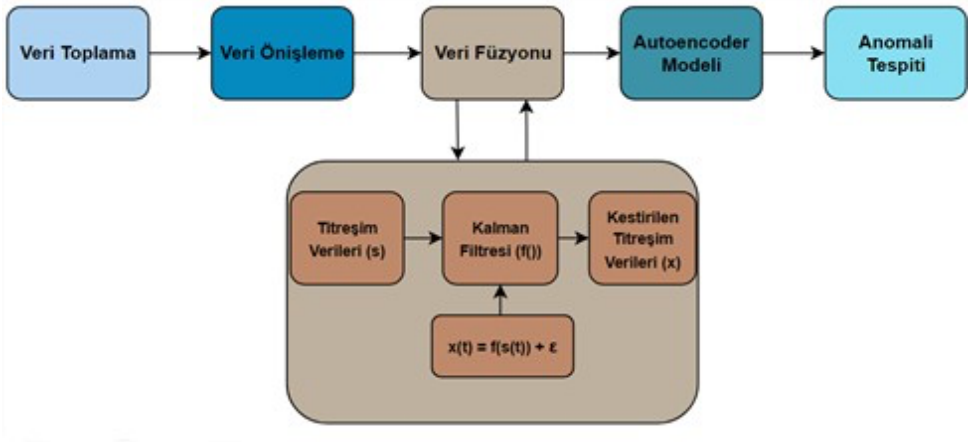
Bu çalışmada geliştirilen yöntemin kuş bakışı bir blok diyagramı Şekil 4'te verilmiştir. Yöntemin blok diyagramında görüldüğü üzere, ilk adımda Şekil 5'te görseli verilen NASA rulman titreşim sinyal verileri girdi olarak alınmıştır. Veriler normalize edilerek bir ön işlemeye tabi tutulduktan sonra veri füzyonu adımına geçilmiştir. Bu adımda, titreşim verileri (s), temel sensör füzyonu bileşenlerinden sensörleri veya verileri temsil etmektedir. Sistem modeli Denklem (6)'daki gibi ifade edilmiştir.

$$x(t) = f(s(t)) + \varepsilon \quad (6)$$

Burada $x(t)$, ilgilenilen parametreleri temsil ederken; ε, gürültüyü ifade etmektedir. Öte yandan $s(t)$, temel sensör füzyonu bileşenlerinden sensörleri veya verileri temsil etmektedir. $f(s(t))$ terimiyle de girdi olarak alınan ön işlenmiş sensör sinyallerinin kalman filtresine tabi tutulduğu temsil edilmektedir. Her bir sinyal verisi kalman filtresi algoritmasına girdi olarak verilmiş ve yeni değerleri tahmin edilmiştir. Bu tahmin edilen yeni sinyal değerleri de AE'yi beslemektedir.

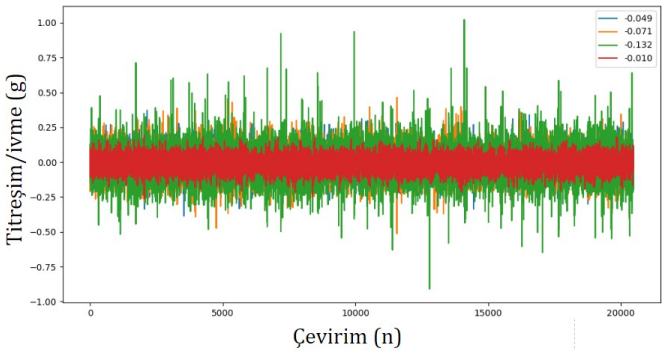
Bir sonraki adımda kalman filtresiyle yeniden kestirimi yapılan sinyaller AE modeline girdi olarak verilmiştir. Veriler, bir zaman serisi verileri olduğu için rastgele eğitim ve test veri kümelerine ayrılmamış ve ilk 445 veri eğitime, geri kalan 538 veri ise teste ayrılmıştır.

Modelin eğitimi aşamasında başarım oranının ölçülmesi için kullanılan metrik kayıp fonksiyonudur. Anomali tespiti uygulamalarında genel olarak ortalama kare hatası (mean squared



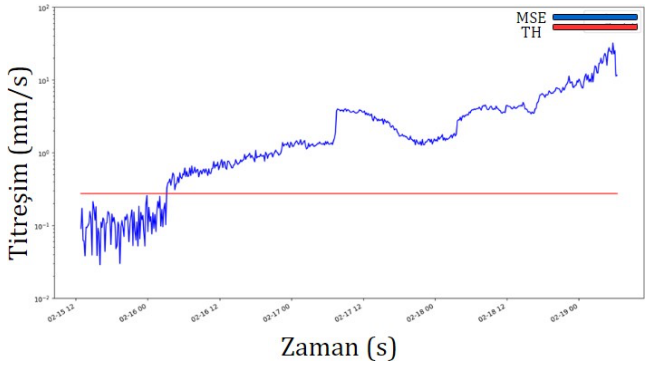
Şekil 4. Sistem blok diyagramı

error, MSE) kullanıldığı için bu çalışmada da MSE tercih edildi. Kayıp fonksiyonunun minimize edilmesi ve diğer model parametrelerinin optimizasyonu için önceden de bahsedilen Adam optimizasyon algoritması çalıştırılmıştır. Algoritmanının öğrenme oranı 0.0001 olarak ayarlanmıştır. Algoritmanın durdurma kriteri olarak da maksimum iterasyon sayısı tercih edilmiş ve 100 olarak belirlenmiştir.



Şekil 5. Rulman titreşim verisi

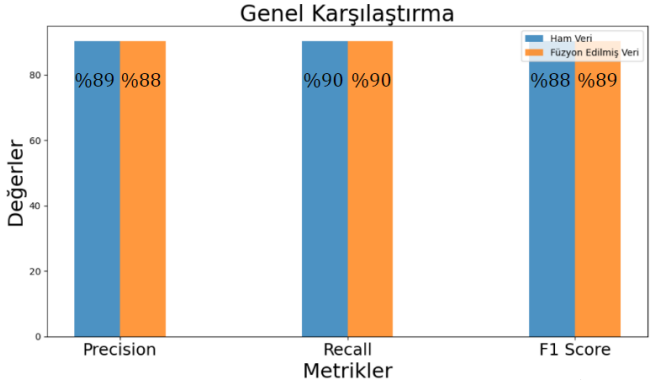
Son adımda, belirlenen parametreler ve durdurma kriteriyle eğitilen AE modeli ile test veri kümesi üzerinde anomali tespiti süreci icra edilmiştir. Bu süreçte, iki farklı deney yürütülmüştür. Yürütülen iki farklı deneyin ilkinde sensör füzyonu adımı pas geçilmiş ve rulman verileri ön işleme tabi tutulduktan sonra direkt olarak AE modeline girdi olarak verilmiştir. İkinci deneyde ise sensör füzyonu adımı sisteme dahil edilmiş ve Kalman filtresiyle ham sinyal verileri üzerinde kestirim yapılmıştır. Uzmanlar tarafından işaretlenen anomali verileri baz alınarak belirlenen eşik değeri kullanılmıştır. İvme büyüklüğü bu eşik değerin üzerinde olan veriler anomali olarak etiketlenmiştir. Şekil 6'da anomali tespiti sürecinin görsel temsili yer almaktadır.



Şekil 6. Anomali tespiti süreci

4 Araştırma bulguları

Deneylerdeki başarım oranlarını ölçebilmek için yaygın olarak kullanılan performans metrikleri seçilmiştir. Bunlar: Doğruluk (Accuracy), Hata (Loss), Kesinlik (Precision), Geri Çağırma (Recall), ROC Eğrisi Altında Kalan Alan (Area Under the Curve Receiver Operating Characteristics) ve F1 Skordur [33]. Elde edilen sonuçlar Şekil 7'de gösterilmiştir. Birinci deneyde sırasıyla kesinlik %89, geri çağırma %90, F1-skor %88; ikinci deneyde sırasıyla kesinlik %88, geri çağırma %90 ve F1-skor %89 olarak hesaplanmıştır.



Şekil 7. Elde edilen metriklerin karşılaştırılması

5 Sonuçlar

Hava araçlarının bakımının ve bileşenlerin sürdürülebilirliğinin sağlanmasında anomali tespiti yaygın kullanılan bir uygulamadır. Günümüzde daha da karmaşık hale gelen havacılık sistemlerinde meydana gelen anomali ve hataların tespiti için geleneksel yöntemler yetersiz kalmaya başlamamıştır. Çoklu sensör füzyonu tekniklerinin icrası ile daha etkin anomali tespiti uygulamaları geliştirilebilmektedir. Bu çalışmada da yaygın olarak kullanılan Kalman filtresi kullanılarak ham sinyal veriler üzerinde gürültülerin azaltılmasıyla yeni tahmin verileri elde edilmiştir. Bu tahmin verileri kullanılarak oldukça yaygın kullanılan AE modeli, tahmin verilerini girdi olarak kullanarak anomali tespiti sürecini icra etmiştir. Bu süreçte parametrelerin ve kayıp fonksiyonunun optimizasyonu için GDA'nın ileri düzey bir versiyonu olan Adam algoritmasına başvurulmuştur. Elde edilen sonuçlara göre kesinlik, geri çağırma ve F1-skorun birbirine çok yakın çıktığı gözlemlenmiştir. Gelecek çalışmalarda farklı kestirim algoritmaları, farklı model hiperparametreleri veya farklı veri kümeleri üzerinde deneyler yapılması planlanmaktadır.

Kaynaklar

[1] Birgelen VA, Buratti D, Mager J, Niggemann O, "Self-organizing maps for anomaly localization and predictive maintenance in cyber-physical production systems". *Procedia CIRP*, 72, 480-485, 2018.

[2] Tian R, Liboni L, Capretz M. "Anomaly Detection with Convolutional Autoencoder for Predictive Maintenance". *9th International Conference on Soft Computing & Machine Intelligence (ISCMI)*, Toronto, ON, Canada, 241-245, 2022.

[3] Alpaydin E. *Introduction to Machine Learning*. Cambridge, MA: MIT Press, 2014.

[4] Nassif AB, Talib MA, Nasir Q, Dakalbab FM. "Machine Learning for Anomaly Detection: A Systematic Review". *IEEE Access*, 9, 78658-78700, 2021.

[5] Denźux T. "40 years of Dempster-Shafer theory". *International Journal of Approximate Reasoning*, 1, 79(C), 1-6, 2016.

[6] Garcia EJ, Beres SL, Mulvihill ML, Stephens CL, Napoli NJ. "Multiclass Flight Anomaly Detection Using Sensor Fusion Based on Dempster-Shafer Theory". *InAIAA Sci. Tech Forum*, 2024.

[7] Bengio Y, Simard P, Frasconi P. "Learning long-term dependencies with gradient descent is difficult". *IEEE Transactions on Neural Networks*, 5(2), 157-66, 1994.

[8] Graves A, Liwicki M, Fernández S, Bertolami R, Bunke H, Schmidhuber J. "A novel connectionist system for unconstrained handwriting recognition". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30, 31(5), 855-68, 2008.

[9] Schmidhuber J, Hochreiter S. "Long short-term memory". *Neural Comput.*, 15, 9(8), 1735-80, 1997.

[10] Korstanje J. "Lstm rnns". *InAdvanced Forecasting with Python: With State-of-the-Art-Models Including LSTMs, Facebook's Prophet, and Amazon's DeepAR,* 243-251, 2021.

[11] Hundman K, Constantinou V, Laporte C, Colwell I, Soderstrom T. "Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding" *24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Jul 19, 387-395, 2018.

[12] Fukushima K. Neocognitron: "A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position". *Biological Cybernetics*, 36(4), 193-202, 1980.

[13] Michelucci U, Michelucci U. "Fundamentals of convolutional neural networks". *Advanced Applied Deep Learning: Convolutional Neural Networks and Object Detection*, 79-123, 2019.

[14] Zeiler MD, Fergus R. "Stochastic pooling for regularization of deep convolutional neural networks". *arXiv preprint*, 1301, 3557, 2013.

[15] Wang M, Wang W, Zhang X, Iu HH. "A new fault diagnosis of rolling bearing based on Markov transition field and CNN". *Entropy*, 24(6):751, 2022.

[16] Qiu H, Lee J, Lin J, Yu G. "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics". *Journal of Sound and Vibration*, 289, (4-5), 1066-90, 2006.

[17] Wang Q, Wang S, Wei B, Chen W, Zhang Y. "Weighted K-NN classification method of bearings fault diagnosis with multi-dimensional sensitive features". *IEEE Access*, 9, 45428-40, 2021.

[18] Liu C, Gryllias K. "A semi-supervised Support Vector Data Description-based fault detection method for rolling element bearings based on cyclic spectral analysis". *Mechanical Systems and Signal Processing*, 140, 106682, 2020.

[19] Ho SK, Nedunuri HC, Balachandran W, Kanfoud J, Gan TH. "Monitoring of industrial machine using a novel blind feature extraction approach". *Applied Sciences*, 22, 11(13), 5792, 2021.

[20] Lee, H. Qiu, G. Yu, J. Lin, and Rexnord Technical Services. IMS, University of Cincinnati. "Bearing Data Set", NASA Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA, 2007.

[21] Kalman, RE. "A new approach to linear filtering and prediction problems.". *J. Basic Eng.*, 82(1), 35-45, 1960.

[22] Welch, G, Bishop, G. "An introduction to the Kalman filter". 1995.

[23] Lee MS, Shifat TA, Hur JW. "Kalman filter assisted deep feature learning for RUL prediction of

hydraulic gear pump". *IEEE Sensors Journal*, 22(11):11088-97, 2022.
[24] Andrychowicz M, Denil M, Gomez S, Hoffman MW, Pfau D, Schaul T, Shillingford B, De Freitas N. "Learning to learn by gradient descent by gradient descent". *Advances in Neural Information Processing Systems*, 29, 2016.
[25] Kinga D, Adam JB. "A method for stochastic optimization". *InInternational Conference on Learning Representations (ICLR)* May 7, 5, 6, 2015.
[26] Kramer, MA. "Nonlinear principal component analysis using autoassociative neural networks". *AIChE Journal*, 37(2), 233-243, 1991.
[27] Wikipedia contributors. "Autoencoder." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia.https://en.wikipedia.org/wiki/Autoe ncoder. Erişim tarihi (3.5.2023).
[28] Nguyen TV, Wong RK, Hegde C. "On the dynamics of gradient descent for autoencoders". *InThe 22nd International Conference on Artificial Intelligence and Statistics*, Apr 11 (2858-2867), PMLR, 2019.
[29] Singh S, Shiv P, Ahmed A. "Outlier modeling in gear bearing using autoencoder for remaining useful life prediction". *Preprints*, 2019.
[30] Ahmad, S, Styp-Rekowski, K, Nedelkoski, S, Kao, O. "Autoencoder-based condition monitoring and anomaly detection method for rotating machines". *In 2020 IEEE International Conference on Big Data (Big Data)*, Atlanta, GA, USA, 4093-4102, 10-13 December, 2020.
[31] Li Z, Sun Y, Yang L, Zhao Z, Chen X. "Unsupervised machine anomaly detection using autoencoder and temporal convolutional network". *IEEE Transactions on Instrumentation and Measurement*, 71, 1-3, 2022.
[32] Raschka S. *Python Machine Learning*. Packt Publishing, 2015.