

**ANKARA ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**YÜKSEK LİSANS TEZİ**

**EVRIŞİMSEL SİNİR AĞLARI TABANLI SU ALTI AKUSTİK HEDEF  
TANIMA**

**Kadir AKBULUT**

**ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**ANKARA  
2026**

**Her hakkı saklıdır**

## ÖZET

Yüksek Lisans Tezi

### EVİRİŞİMSEL SİNİR AĞLARI TABANLI SU ALTI AKUSTİK HEDEF TANIMA

Kadir AKBULUT

Ankara Üniversitesi  
Fen Bilimleri Enstitüsü  
Elektrik-Elektronik Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Asım Egemen YILMAZ

Su altı akustik hedef tanıma geleneksel olarak sonar operatörlerinin düşük frekans analizi ve kaydı (LOFAR – Low Frequency Analysis and Recording) ve gürültü zarf modülasyonu tespiti (DEMON – Detection of Envelope Modulation on Noise) spektrumlarını yorumlamasına dayanır. Ancak operatör başarımının psikofizyolojik etkenlere duyarlılığı bu yaklaşımı sınırlar. Bu operatör bağımlılığını azaltmak amacıyla, spektrogram girdileriyle çalışan evrişimsel sinir ağları tabanlı otomatik bir sınıflandırma sistemi geliştirilmiştir. Başarım üzerinde etkili olan iki tasarım değişkeni olan segment uzunluğu ve Gauss gürültüsü ekleme rejimi, faktöriyel bir deney tasarımı kapsamında birlikte incelenmiştir. ShipsEar veri seti beş sınıfta (dört gemi sınıfı ve ortam gürültüsü) gruplandırılmış; kayıtlar 16 kHz'e yeniden örneklenip kareler ortalamasının karekökü normalizasyonundan geçirilmiş, kısa zamanlı Fourier dönüşümü ile spektrogramlara dönüştürülmüş ve sabit global desibel değerleriyle ölçeklenmiştir. Segment uzunluğu dört düzeyde (10, 5, 3 ve 1 saniye), gürültü rejimi iki düzeyde [ $N(0, 5^2)$ ,  $N(3, 5^2)$ ] ele alınarak sekiz ana deney ve gürültü eklenmemiş 5 saniyelik bir kontrol deneyi oluşturulmuş; ImageNet üzerinde önceden eğitilmiş EfficientNetB1 modeli transfer öğrenmeyle uyarlanıp tabakalı 10 katlı çapraz doğrulama ile değerlendirilmiştir. Doğruluk segment uzunluğuyla monoton olarak değişmemiş, her iki rejimde de 3 saniyelik segment uzunluğunda en yüksek değere ulaşmıştır. İki rejim arasındaki fark tüm segment uzunluklarında 7-8 yüzde puanı düzeyinde sabit kalmış; bu durum, segment uzunluğu ile gürültü rejiminin başarım üzerindeki etkilerinin birbirinden bağımsız olduğuna işaret etmektedir. 5 saniyelik segment uzunluğunda, gürültü eklenmemiş kontrole (%82,37) kıyasla  $N(0, 5^2)$  rejimi doğruluğu %86,69'a yükseltmiş,  $N(3, 5^2)$  rejimi ise kontrol değerinin altına, %79,67'ye düşürmüştür. En yüksek başarım (%88,52  $\pm$  %1,39; makro F1 0,8887) 3 saniyelik segment uzunluğunda ve  $N(0, 5^2)$  rejiminde elde edilmiştir. Bulgular, segment uzunluğunun yalnızca bir veri artırma parametresi değil, aynı zamanda fiziksel olarak anlamlı akustik bağlamı koruyan bir optimizasyon parametresi olduğunu göstermektedir.

**Haziran 2026, 79 sayfa**

**Anahtar Kelimeler:** Su altı akustik hedef tanıma, Segment uzunluğu, Evrişimsel sinir ağları, Sonar, Spektrogram.

## ABSTRACT

Master Thesis

### UNDERWATER ACOUSTIC TARGET RECOGNITION BASED ON CONVOLUTIONAL NEURAL NETWORKS

Kadir AKBULUT

Ankara University  
Graduate School of Natural and Applied Science  
Department of Electrical and Electronics Engineering

Supervisor: Prof. Dr. Asım Egemen YILMAZ

Underwater acoustic target recognition traditionally relies on sonar operators interpreting low frequency analysis and recording (LOFAR) and detection of envelope modulation on noise (DEMON) spectra. However, the sensitivity of operator performance to psychophysiological factors limits this approach. In order to reduce this operator dependency, an automatic classification system based on convolutional neural networks (CNNs) operating on spectrogram inputs was developed. Two design variables affecting performance, namely segment length and the Gaussian noise addition regime, were examined together within a factorial experimental design. The ShipsEar dataset was grouped into five classes (four vessel classes and ambient noise); recordings were resampled to 16 kHz, normalized by root mean square (RMS), transformed into spectrograms using short-time Fourier transform (STFT), and scaled with fixed global decibel values. Segment length at four levels (10, 5, 3, and 1 s) and noise regime at two levels  $[N(0, 5^2), N(3, 5^2)]$  produced eight main experiments plus a 5-second control experiment without added noise; an EfficientNetB1 model pretrained on ImageNet was adapted by transfer learning and evaluated under stratified 10-fold cross-validation. Accuracy varied non-monotonically with segment length, peaking at the 3-second segment length under both regimes. The difference between the two regimes remained constant at a level of 7-8 percentage points across all segment lengths, indicating that the effects of segment length and noise regime on performance are independent of one another. At the 5-second segment length, relative to the control without added noise (82.37%), the  $N(0, 5^2)$  regime raised accuracy to 86.69%, whereas the  $N(3, 5^2)$  regime lowered it below the control value, to 79.67%. The highest performance ( $88.52\% \pm 1.39\%$ ; macro F1 0.8887) was obtained at the 3-second segment length under the  $N(0, 5^2)$  regime. These findings show that segment length is not merely a data-augmentation parameter but an optimization parameter that preserves physically meaningful acoustic context.

**June 2026, 79 pages**

**Key Words:** Underwater acoustic target recognition, Segment length, Convolutional neural networks, Sonar, Spectrogram.

## **TEŞEKKÜR**

Bu tezin gerekleřtirilmesinde, kıymetli grüş ve katkılarıyla beni yönlendiren değerli danışman hocam Prof. Dr. Asım Egemen YILMAZ'a teşekkürlerimi sunarım.

Kadir AKBULUT  
Ankara, Haziran 2026

## İÇİNDEKİLER

|                                                                                      |      |
|--------------------------------------------------------------------------------------|------|
| TEZ ONAY SAYFASI                                                                     |      |
| ETİK.....                                                                            | i    |
| ÖZET.....                                                                            | ii   |
| ABSTRACT .....                                                                       | iii  |
| TEŞEKKÜR .....                                                                       | iv   |
| SİMGELER DİZİNİ .....                                                                | vii  |
| ŞEKİLLER DİZİNİ .....                                                                | viii |
| ÇİZELGELER DİZİNİ .....                                                              | xi   |
| 1. GİRİŞ .....                                                                       | 1    |
| 2. KURAMSAL TEMELLER.....                                                            | 5    |
| 2.1 Sonarın Tarihsel Gelişimi .....                                                  | 5    |
| 2.2 Su Altı Hedeflerinin Tespiti için Neden Radar Kullanılmaz? .....                 | 9    |
| 2.3 Zaman-Frekans Ayırıştırma Yöntemleri .....                                       | 10   |
| 2.3.1 Kısa zamanlı Fourier dönüşümü (Short-time Fourier transform).....              | 11   |
| 2.3.2 Dalgacık dönüşümü (Wavelet transform).....                                     | 11   |
| 2.3.3 Mel-frekans kepsral katsayıları (Mel-frequency cepstral coefficients) .....    | 11   |
| 2.4 Spektrogram .....                                                                | 12   |
| 3. SONAR TEKNOLOJİSİNİN UYGULAMA ALANLARI VE<br>DENİZCİLİK SEKTÖRÜNDEKİ KONUMU ..... | 14   |
| 3.1 Sonar Teknolojisinin Uygulama Alanları .....                                     | 14   |
| 3.1.1 Askeri uygulamalar.....                                                        | 14   |
| 3.1.1.1 Denizaltı savunma harbi.....                                                 | 15   |
| 3.1.1.2 Mayın tespiti ve karşı tedbirler.....                                        | 15   |
| 3.1.1.3 Liman ve kritik tesis güvenliği.....                                         | 16   |
| 3.1.1.4 Su altı gözetleme ve keşif .....                                             | 17   |
| 3.1.1.5 Torpidoya karşı tedbirler .....                                              | 17   |
| 3.1.1.6 Denizaltı eğitimi.....                                                       | 17   |
| 3.1.2 Sivil uygulamalar .....                                                        | 17   |
| 3.1.2.1 Balıkçılık ve su ürünleri yetiştiriciliği .....                              | 18   |
| 3.1.2.2 Oşinografik araştırmalar ve deniz tabanı haritalaması .....                  | 18   |
| 3.1.2.3 Petrol ve gaz keşfi.....                                                     | 19   |
| 3.1.2.4 Arama ve kurtarma operasyonları.....                                         | 19   |
| 3.1.2.5 Su altı altyapı denetimi ve bakımı .....                                     | 19   |
| 3.1.2.6 Tıbbi uygulamalar.....                                                       | 19   |
| 3.1.2.7 Endüstriyel izleme ve tahmini bakım .....                                    | 20   |
| 3.2 Denizcilik Sektöründe Kullanılan Sonar Konfigürasyonları .....                   | 20   |
| 3.2.1 Karinaya monteli sonar .....                                                   | 20   |
| 3.2.2 Çekili dizin sonar (TAS-Towed Array Sonar) .....                               | 20   |
| 3.2.3 Değişken derinlikli sonar (VDS-Variable Depth Sonar) .....                     | 21   |
| 3.2.4 Daldırma sonarı.....                                                           | 21   |
| 3.2.5 Yandan taramalı sonar .....                                                    | 21   |
| 3.2.6 Sentetik açıklıklı sonar (SAS-Synthetic Aperture Sonar) .....                  | 21   |
| 3.2.7 İskandil (Echo Sounder).....                                                   | 22   |
| 4. MATERYAL VE YÖNTEM.....                                                           | 23   |
| 4.1 Giriş ve Problem Tanımı .....                                                    | 23   |

|                                                                                                                  |    |
|------------------------------------------------------------------------------------------------------------------|----|
| 4.2 Veri Seti ve Ön İşleme .....                                                                                 | 25 |
| 4.3 Spektrogram Oluşturma ve Sınıf Dağılımları .....                                                             | 32 |
| 4.4 CNN Model Mimarisi ve Eğitim Detayları .....                                                                 | 32 |
| 4.5 Performans Değerlendirme Kriterleri .....                                                                    | 35 |
| 4.5.1 Hata matrisi .....                                                                                         | 36 |
| 4.5.2 Performans metrikleri .....                                                                                | 36 |
| 4.6 Deneysel Sonuçlar ve Performans Analizi .....                                                                | 38 |
| 4.6.1 10 saniyelik segmentasyon deneyi sonuçları .....                                                           | 40 |
| 4.6.2 5 saniyelik segmentasyon deneyi sonuçları .....                                                            | 47 |
| 4.6.3 3 saniyelik segmentasyon deneyi sonuçları .....                                                            | 54 |
| 4.6.4 1 saniyelik segmentasyon deneyi sonuçları .....                                                            | 60 |
| 4.6.5 Segment uzunluğu ve gürültü rejiminin model performansına etkileri<br>üzerine karşılaştırmalı analiz ..... | 66 |
| 5. SONUÇ VE ÖNERİLER .....                                                                                       | 70 |
| KAYNAKLAR .....                                                                                                  | 76 |
| ÖZGEÇMİŞ .....                                                                                                   | 79 |

## SİMGELER DİZİNİ

|                    |                                                                      |
|--------------------|----------------------------------------------------------------------|
| dB                 | : Desibel                                                            |
| Hz                 | : Hertz                                                              |
| kHz                | : Kilohertz                                                          |
| MHz                | : Megahertz                                                          |
| log                | : Logaritma                                                          |
| °C                 | : Santigrat Derece                                                   |
| m                  | : Metre                                                              |
| $N(\mu, \sigma^2)$ | : Ortalaması $\mu$ , varyansı $\sigma^2$ olan normal (Gauss) dağılım |

## Kısaltmalar

|       |                                                   |
|-------|---------------------------------------------------|
| ABD   | Amerika Birleşik Devletleri                       |
| AUV   | Autonomous Underwater Vehicle                     |
| BT    | Batitermograf                                     |
| CNN   | Convolutional Neural Network                      |
| DCGAN | Deep Convolutional Generative Adversarial Network |
| DEMON | Detection of Envelope Modulation on Noise         |
| DN    | Doğru Negatif                                     |
| DP    | Doğru Pozitif                                     |
| DSH   | Denizaltı Savunma Harbi                           |
| DTS   | Dalgıç Tespit Sonarı                              |
| EEG   | Elektroensefalogram                               |
| EKG   | Elektrokardiyogram                                |
| LOFAR | Low Frequency Analysis and Recording              |
| MFCC  | Mel-Frequency Cepstral Coefficients               |
| MIT   | Massachusetts Institute of Technology             |
| Radar | Radio Detection and Ranging                       |
| ReLU  | Rectified Linear Unit                             |
| RMS   | Root Mean Square                                  |
| ROV   | Remotely Operated Vehicle                         |
| SAS   | Sentetik Açıklıklı Sonar                          |
| SNR   | Signal-to-Noise Ratio                             |
| SOFAR | Sound Fixing and Ranging                          |
| Sonar | Sound Navigation and Ranging                      |
| SOSUS | Sound Surveillance System                         |
| STFT  | Short-Time Fourier Transform                      |
| UUV   | Unmanned Underwater Vehicle                       |
| VDS   | Variable Depth Sonar                              |
| WHOI  | Woods Hole Oceanographic Institution              |
| YN    | Yanlış Negatif                                    |
| YP    | Yanlış Pozitif                                    |

## ŞEKİLLER DİZİNİ

|                                                                                                                 |    |
|-----------------------------------------------------------------------------------------------------------------|----|
| Şekil 2.1 Charles Sturm (soldaki) ve Daniel Colladon (sağdaki), 1826'da sesin sudaki ilk doğru ölçümü.....      | 6  |
| Şekil 2.2 Sıcaklık değişimlerine bağlı sesin kırılma ve bükülmesinden dolayı oluşan gölge bölge.....            | 7  |
| Şekil 2.3 SOFAR kanalı.....                                                                                     | 8  |
| Şekil 2.4 Su altında neden radar kullanılmaz? .....                                                             | 9  |
| Şekil 3.1 Nantucket kıyılarında kimliği belirlenemeyen bir gemi enkazının SAS taraması.....                     | 16 |
| Şekil 3.2 Yandan taramalı sonar ve sentetik açıklıklı sonar çalışma yapısının karşılaştırılması .....           | 16 |
| Şekil 4.1 Çalışmada izlenen metodolojinin genel akış şeması .....                                               | 25 |
| Şekil 4.2 Class A'ya ait örnek spektrogramlar.....                                                              | 27 |
| Şekil 4.3 Class B'ye ait örnek spektrogramlar.....                                                              | 27 |
| Şekil 4.4 Class C'ye ait örnek spektrogramlar.....                                                              | 28 |
| Şekil 4.5 Class D'ye ait örnek spektrogramlar.....                                                              | 28 |
| Şekil 4.6 Class E'ye ait örnek spektrogramlar .....                                                             | 29 |
| Şekil 4.7 EfficientNetB1 mimarisi.....                                                                          | 33 |
| Şekil 4.8 EfficientNetB1 modelinin parametreleri .....                                                          | 33 |
| Şekil 4.9 10 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu eğitim ve doğrulama doğrulukları .....  | 41 |
| Şekil 4.10 10 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu eğitim ve doğrulama kayıpları .....    | 41 |
| Şekil 4.11 10 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu hata matrisi .....                     | 42 |
| Şekil 4.12 10 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu normalize hata matrisi .....           | 43 |
| Şekil 4.13 10 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu eğitim ve doğrulama doğrulukları ..... | 44 |
| Şekil 4.14 10 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu eğitim ve doğrulama kayıpları .....    | 45 |
| Şekil 4.15 10 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu hata matrisi.....                      | 46 |
| Şekil 4.16 10 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu normalize hata matrisi .....           | 46 |
| Şekil 4.17 5 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu eğitim ve doğrulama doğrulukları .....  | 48 |

|                                                                                                                |    |
|----------------------------------------------------------------------------------------------------------------|----|
| Şekil 4.18 5 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu eğitim ve doğrulama kayıpları .....    | 48 |
| Şekil 4.19 5 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu hata matrisi .....                     | 49 |
| Şekil 4.20 5 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu normalize hata matrisi .....           | 49 |
| Şekil 4.21 5 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu eğitim ve doğrulama doğrulukları ..... | 51 |
| Şekil 4.22 5 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu eğitim ve doğrulama kayıpları .....    | 51 |
| Şekil 4.23 5 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu hata matrisi.....                      | 52 |
| Şekil 4.24 5 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu normalize hata matrisi .....           | 52 |
| Şekil 4.25 3 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu eğitim ve doğrulama doğrulukları ..... | 55 |
| Şekil 4.26 3 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu eğitim ve doğrulama kayıpları .....    | 55 |
| Şekil 4.27 3 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu hata matrisi.....                      | 56 |
| Şekil 4.28 3 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu normalize hata matrisi .....           | 56 |
| Şekil 4.29 3 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu eğitim ve doğrulama doğrulukları ..... | 58 |
| Şekil 4.30 3 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu eğitim ve doğrulama kayıpları .....    | 58 |
| Şekil 4.31 3 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu hata matrisi .....                     | 59 |
| Şekil 4.32 3 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu normalize hata matrisi .....           | 59 |
| Şekil 4.33 1 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu eğitim ve doğrulama doğrulukları ..... | 61 |
| Şekil 4.34 1 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu eğitim ve doğrulama kayıpları .....    | 61 |
| Şekil 4.35 1 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu hata matrisi .....                     | 62 |
| Şekil 4.36 1 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu normalize hata matrisi .....           | 62 |
| Şekil 4.37 1 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu eğitim ve doğrulama doğrulukları ..... | 64 |

|                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------|----|
| Şekil 4.38 1 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu eğitim ve doğrulama kayıpları ..... | 64 |
| Şekil 4.39 1 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu hata matrisi.....                   | 65 |
| Şekil 4.40 1 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu normalize hata matrisi .....        | 65 |

## ÇİZELGELER DİZİNİ

|                                                                                                          |    |
|----------------------------------------------------------------------------------------------------------|----|
| Çizelge 4.1 Oluşturulan 5 sınıfın ayrıntılı açıklaması.....                                              | 25 |
| Çizelge 4.2 Sekiz ana deneyin ve gürültüsüz kontrol deneyinin özet başarımlar<br>göstergeleri .....      | 39 |
| Çizelge 4.3 10 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu performans<br>metrikleri ..... | 43 |
| Çizelge 4.4 10 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu performans<br>metrikleri ..... | 47 |
| Çizelge 4.5 5 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu performans<br>metrikleri .....  | 50 |
| Çizelge 4.6 5 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu performans<br>metrikleri .....  | 53 |
| Çizelge 4.7 5 saniyelik segmentasyon deneyi (gürültü eklenmeden) sonucu<br>performans metrikleri .....   | 54 |
| Çizelge 4.8 3 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu performans<br>metrikleri .....  | 57 |
| Çizelge 4.9 3 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu performans<br>metrikleri .....  | 60 |
| Çizelge 4.10 1 saniyelik segmentasyon deneyi ( $N(0, 5^2)$ rejimi) sonucu performans<br>metrikleri ..... | 63 |
| Çizelge 4.11 1 saniyelik segmentasyon deneyi ( $N(3, 5^2)$ rejimi) sonucu performans<br>metrikleri ..... | 66 |

## 1. GİRİŞ

Su altı ortamı, deniz suyunun yüksek elektriksel iletkenliği nedeniyle elektromanyetik dalgaların onlarca metre mesafede üstel olarak soğurulduğu, akustik dalgaların ise yüzlerce kilometre uzaklığa görece düşük bir yayılım kaybıyla iletilebildiği kendine özgü bir fiziksel ortamdır. Elektromanyetik ve akustik yayılım arasındaki bu temel fiziksel farklılık, su altında hedef tespiti, sınıflandırılması, yerinin belirlenmesi ve takip edilmesi için sonar (Sound Navigation and Ranging) teknolojisini birincil algılama yöntemi olarak konumlandırmıştır (Kaharlı 1999, Neupane ve Seok 2020). 19. yüzyılda Colladon ve Sturm'un Cenevre Gölü'nde sesin sudaki hızını ilk kez doğrulukla ölçmesinden II. Dünya Savaşı'nın ardından okyanus tabanına döşenen SOSUS (Sound Surveillance System) hidrofon dizilerine ve günümüzde otonom su altı araçlarına entegre edilen çok ışınlı iskandillere uzanan yaklaşık iki yüzyıllık birikim, sonarı modern denizciliğin temel algılama altyapısı hâline getirmiştir (Kaharlı 1999).

Sonar teknolojileri, su altı ortamlarında hedef tespiti ve sınıflandırması için kritik bir rol oynamaktadır. Denizaltı savunma harbi (DSH), mayın avcılığı ve kritik deniz altyapısının korunması gibi askeri uygulamaların yanı sıra; oşinografik araştırmalar, deniz tabanı haritalama, hidrokarbon arama, arama-kurtarma operasyonları ve su altı altyapı denetimi gibi sivil uygulamalar bu teknolojinin geniş kullanım yelpazesini oluşturmaktadır.

Geleneksel pasif sonar sınıflandırma yaklaşımları LOFAR ile DEMON spektrumlarının operatörler tarafından görsel ve işitsel olarak yorumlanmasına dayanmaktadır (Aslam vd. 2024, Neupane ve Seok 2020). Bu paradigma çeşitli yapısal sınırlılıklar içermektedir. İlk olarak yorumlama performansı operatörün deneyimi, dikkati, fiziksel ve psikolojik durumu ile doğrudan ilişkilidir ve uzun süreli görevlerde dalgalanmaya açıktır. İkinci olarak deneyimli sonar operatörü yetiştirilmesi zaman ve maliyet açısından yüksek bir yatırımı gerektirmekte, nitelikli personel kaynağı sınırlı kalmaktadır. Üçüncü olarak su altı akustik ortamı; biyolojik gürültü, doğal gürültü (rüzgâr, yağmur, dalga), antropojenik gürültü (ticari deniz trafiği, sismik araştırmalar) ve çoklu yol yayılımı içeren karmaşık bir akustik ortama sahiptir (Aslam vd. 2024). Son olarak gemilerden yayılan gürültü; durağan olmayan, doğrusal olmayan ve geniş bir dinamik aralıkta değişen karakteristikler

sergilediği için geleneksel doğrusal sinyal işleme teknikleri sıklıkla yetersiz kalmaktadır (Aslam vd. 2024).

Bu sınırlılıkların aşılması, operatör bağımlılığını azaltan ve sınıflandırma kararlarını daha tutarlı ve hızlı üreten otomatik su altı akustik hedef tanıma sistemlerinin geliştirilmesini zorunlu kılmaktadır.

Akustik sinyaller doğaları gereği zamanla değişen spektral içerik taşıdığından, yalnızca zaman bölgesinde veya yalnızca frekans bölgesinde elde edilecek temsiller bilgi kaybına neden olmaktadır. Bu nedenle STFT, dalgacık dönüşümü, mel-frekans kepsstral katsayıları (MFCC) ve spektrogram gibi zaman-frekans temsilleri, sonar sinyallerinin analizinde merkezi bir rol üstlenmiştir (Aslam vd. 2024, Neupane ve Seok 2020). Bu temsiller arasında spektrogram; iki boyutlu görsel yapısı ve hem dar bantlı tonal bileşenleri hem de geniş bantlı stokastik bileşenleri eş zamanlı barındırma kapasitesi nedeniyle özellikle CNN tabanlı sınıflandırıcılar için doğal bir girdi formatı oluşturmaktadır (Dörfler vd. 2017, Zhu vd. 2022).

Su altı akustik hedef tanıma alanında derin öğrenme tabanlı yaklaşımlara geçiş, son yıllarda büyük bir ivme kazanmıştır (Aslam vd. 2024, Neupane ve Seok 2020). Hong vd. (2021), optimize edilmiş öznitelik çıkarımı yaklaşımıyla ResNet18 mimarisini kullanarak ShipsEar veri setinde %94,3 doğruluk oranı elde etmiştir. Sinyal dönüşüm adımlarını azaltmaya yönelik çalışmalarda, Shen vd. (2018) ham hidrofona verisinden doğrudan öğrenim sağlayan işitsel tabanlı bir CNN tanıtmıştır. Benzer şekilde Doan vd. (2022) ise zaman bölgesindeki ham ses sinyalini, herhangi bir spektrogram dönüşümü uygulamaksızın doğrudan girdi olarak kabul eden yoğun bağlantılı bir CNN mimarisi önermiştir. Spektrogram tabanlı güncel yaklaşımlar incelendiğinde; Liu vd. (2021) üç boyutlu Mel-spektrogram özniteliklerini, veri çoğaltma tekniklerini ve evrimsel-tekrarlayan sinir ağını (CRNN) bütünleştiren bir yöntem sunmuştur. Yang vd. (2024) tarafından geliştirilen, alt-bant birleştirilmiş Mel-spektrogram (SC-Mel) ve çoklu-bölge dikkat mekanizması tabanlı CFTANet modeli ise DeepShip ve ShipsEar veri setlerinde sınanmıştır. Söz konusu model, DeepShip veri setinde %90,60 doğruluk ile literatürdeki en gelişmiş başarımlar (state-of-the-art) seviyesine ulaşırken, ShipsEar veri setinde %96,40

doğruluk sergilemiştir. Öznitelik ve veri zenginleştirme odaklı diğer çalışmalardan Yao vd. (2023), Mel-frekans kepsral katsayıları (MFCC) özniteliklerini ve DCGAN (Deep Convolutional Generative Adversarial Network) tabanlı veri çoğaltmayı artık (residual) bağlantılı bir CNN (ResNet18) mimarisiyle birleştirmiştir. Mimarinin kompaktlığına odaklanan Zeng vd. (2025), kritik frekans bandına yoğunlaşan ve frekans kanallaştırması uygulayan hafif bir yapı (FCResNet5) önerirken; Xu vd. (2023) ise pürüzsüzlük sağlayan düzenleme ile yerel maskeleye ve kopyalama (LMR) tabanlı bir spektrogram çoğaltma stratejisi geliştirmiştir.

Bu tez çalışması, ShipsEar veri seti üzerinde spektrogram tabanlı bir CNN kullanarak su altı akustik hedeflerin otomatik sınıflandırılması için bütüncül ve yeniden üretilebilir bir metodoloji geliştirmeyi; ayrıca segment uzunluğu ile eğitim setine uygulanan Gauss gürültüsü artırımı rejiminin model performansı üzerindeki etkilerini faktöriyel bir tasarım çerçevesinde sistematik biçimde incelemeyi amaçlamaktadır. Bu yaklaşım, sonar operatörlerinin üzerindeki yükü azaltarak operasyonel etkinliği artırmayı ve insan kaynaklı hataları en aza indirmeyi hedeflemektedir. Literatürde, yapay zekâ destekli hedef tespit ve sınıflandırma üzerine yapılan çalışmalar, bu alandaki potansiyeli göstermektedir.

Bu tezin ikinci bölümünde, sonar teknolojisinin tarihsel gelişimi ve temel kuramsal temelleri detaylı bir şekilde ele alınmıştır. Bu kısımda, sesin su altındaki yayılım özellikleri, aktif ve pasif sonar sistemlerinin çalışma prensipleri açıklanmış; ayrıca su altında radar yerine neden sonar kullanıldığı deniz suyunun yüksek iletkenliği nedeniyle yaşanan elektromanyetik kayıpların büyüklüğü üzerinden ele alınmaktadır. Bu bölümde ayrıca, sinyal işlemede kritik rol oynayan zaman-frekans ayrıştırma yöntemleri olan STFT, Dalgacık Dönüşümü (Wavelet Transform), Mel-Frekans Kepsral Katsayıları (MFCC) ve spektrogram gibi kavramlara değinilmiştir.

Tezin üçüncü bölümünde, sonar teknolojisinin küresel çaptaki uygulama alanları, denizcilik sektöründeki konumu kapsamlı bir şekilde sunulmuştur. Bu bölümde, sonar teknolojisinin askeri uygulamaları (Denizaltı Savunma Harbi, Mayın Tespiti, Liman ve Kritik Tesis Güvenliği, Su Altı Gözetleme, Torpidoya Karşı Tedbirler, Denizaltı Eğitimi)

ile sivil uygulamaları (Balıkçılık, Oşinografik Araştırmalar, Petrol ve Gaz Keşfi, Arama Kurtarma Operasyonları, Su Altı Altyapı Denetimi, Tıbbi Uygulamalar, Endüstriyel İzleme) detaylı olarak açıklanmıştır.

Tezin dördüncü bölümünde, CNN tabanlı hedef sınıflandırma sisteminin metodolojisi ve gerçekleştirilen deneysel çalışmalar ayrıntılı biçimde sunulmaktadır. Bu kapsamda ön işleme zinciri, dört farklı segmentasyon senaryosu altında üretilen spektrogramların eğitime hazırlanması, EfficientNetB1 mimarisinin transfer öğrenme yapılandırması ve veri seti ayırma oranları, epoch sayısı, erken durdurma, öğrenme oranı azaltma ile en iyi modeli kaydetme gibi geri çağırma fonksiyonlarını içeren eğitim hiperparametreleri adım adım açıklanmıştır. Deneysel çalışmalar, segment uzunluğunun yanı sıra eğitim setine uygulanan Gauss gürültüsü artırımının da bir tasarım değişkeni olarak ele alındığı faktöriyel bir yapıda kurgulanmış; bu kapsamda gürültü rejimi,  $N(0, 5^2)$  ve  $N(3, 5^2)$  olmak üzere iki düzeyde incelenerek dört segment uzunluğu ile birlikte sekiz ana deney oluşturulmuştur. Bu deneylere ek olarak, gürültü artırımının başarıma net katkısını yalıtılabilmek amacıyla 5 saniyelik segmentasyon için gürültü uygulanmayan bir kontrol deneyi de yürütülmüştür. Geliştirilen sistemin performans değerlendirmesi için kullanılan hata matrisi (confusion matrix) ile kesinlik (precision), duyarlılık (recall) ve F1-Skoru metrikleri tanımlanmış olup bu metrikler üzerinden söz konusu deney konfigürasyonları karşılaştırmalı olarak analiz edilerek modelin genelleme yeteneğini en üst düzeye çıkaran segment uzunluğu ve gürültü rejimi birlikteliği deneysel olarak belirlenmiştir.

Sonuç bölümünde ise segment uzunluğu ve gürültü rejiminin sınıflandırma başarımı üzerindeki etkileri temelinde elde edilen deneysel bulgular sistematik olarak özetlenmekte; en yüksek başarıma sağlayan segment uzunluğu ve gürültü rejimi birlikteliğine ilişkin nicel kanıtlar, gürültü artırımının çift yönlü etkisi, sınıf bazlı performans farklılıkları ve modelin genelleme kapasitesi değerlendirilmektedir. Bölüm ayrıca, operatör bağımlılığının azaltılması, yeniden üretilebilir bir metodolojinin sağlanması ve gemi akustik imzasının fiziksel doğasıyla uyumlu (özgün akustik karakteristiği koruyan) veri çoğaltma stratejilerinin geliştirilmesi açılarından çalışmanın katkılarını değerlendirmektedir.

## 2. KURAMSAL TEMELLER

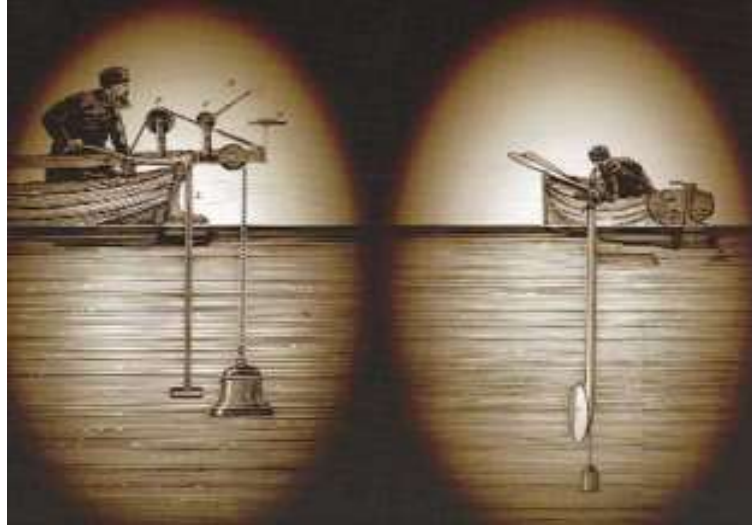
### 2.1 Sonarın Tarihsel Gelişimi

Sonar (Sound Navigation and Ranging) su altı hedeflerinin tespiti, sınıflandırılması, yerinin belirlenmesi, tanımlanması ve izlenmesi için kullanılır. Sonar sistemleri temelde aktif ve pasif sonar olmak üzere ikiye ayrılır (Neupane ve Seok 2020, Waite 2002). Aktif sonar, transdüser kullanarak elektrik sinyalini akustik sinyale dönüştürür. Su ortamında ilerleyen akustik sinyal bir engelle çarptığında yansıma meydana gelir. Transdüser yansıyan akustik sinyalleri tekrar elektrik sinyaline dönüştürür. Bu elektrik sinyali aktif sonar tarafından algılanır ve analiz edilir. Analiz sonucunda hedefin tipi, uzaklığı ve yönü hesaplanır. Pasif sonar bir su altı mikrofonu olan hidrofonu kullanarak su altı nesnelerinden (denizaltılar veya gemiler gibi) ve balinalar gibi deniz hayvanlarından gelen gürültüyü tespit etmek için kullanılır. Aktif sonarın aksine pasif sonar kendi sinyalini yaymaz. Bu, tespit edilmek istemeyen askeri gemiler veya okyanusu sessizce dinlemeye odaklanan bilimsel görevler için bir avantajdır. Pasif sonar kendi sinyalini yaymak yerine yalnızca kendisine doğru gelen ses dalgalarını tespit eder. Pasif sonar diğer pasif dinleme cihazlarıyla birlikte kullanılmadığı sürece bir nesnenin menzilini ölçemez. Çoklu (en az üç) pasif sonar cihazı, bir ses kaynağının üçgenlenmesine izin verebilir (National Oceanic and Atmospheric Administration 2024b, Waite 2002).

Görevin ihtiyaçlarına bağlı olarak, yüksek ve düşük frekansların kullanımı değiştirilebilir. Daha yüksek ses frekansları daha iyi görüntü çözünürlüğü elde etmek için daha düşük ses frekansları ise uzun mesafeler kat etmek için kullanılır (Neupane ve Seok 2020).

Doğada bazı hayvanların (yunus, yarası ve diğerleri) haberleşme ve nesne tespiti için ses kullandığı bilinse de sesin insanlar tarafından su altında kullanılmaya başlanması 1490 yılında Leonardo da Vinci ile olmuştur. “Gemini durdurup uzun bir borunun ucunu suya daldırır ve öteki ucunu kulağına dayarsan, senden çok uzakta bulunan gemilerin seslerini işitebilirsin.” (Leonardo da Vinci, 1490). 1687’de ses yayılımının ilk matematiksel teorisi Sir Isaac Newton’ın *Philosophiae Naturalis Principia Mathematica* adlı eserinde yayınlanmıştır. Sesin havadaki hızı 17. yüzyılın ortalarından itibaren ölçülmeye başlandı

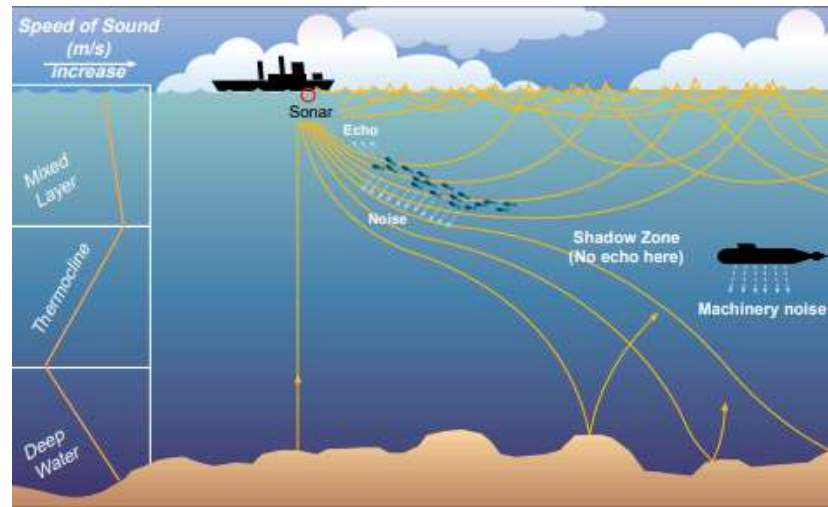
ancak İsviçreli fizikçi Daniel Colladon ve Fransız matematikçi Charles Sturm sesin sudaki hızını doğru bir şekilde 1826'da ölçebildiler. Cenevre Gölü'nde yaptıkları bu ölçümde su altına bir zil daldırdılar ve bu zilden yayılan sesi dinleyerek ses hızını ölçmeye çalıştılar. Söz konusu çalışmada uzun bir boru kullanarak dinleme yaptılar (Leonardo da Vinci'nin önerdiği gibi) ve sesin ne kadar sürede belirlenen mesafeye ulaştığını kronometre ile ölçerek ses hızını ölçebildiler. Sesin hızını su sıcaklığı  $1,8^{\circ}\text{C}$ 'de saniyede 1435 metre olarak ölçtüler ve bu ölçüm bugün kabul edilen hızdan sadece saniyede 3 metre uzaktaydı. Bu araştırmacıların gösterdiği şey, suyun ses için mükemmel bir ortam olduğu ve sesi havadaki hızından neredeyse 5 kat daha hızlı iletmişti (Kaharl 1999).



Şekil 2.1 Charles Sturm (soldaki) ve Daniel Colladon (sağdaki), 1826'da sesin sudaki ilk doğru ölçümü (Kaharl 1999)

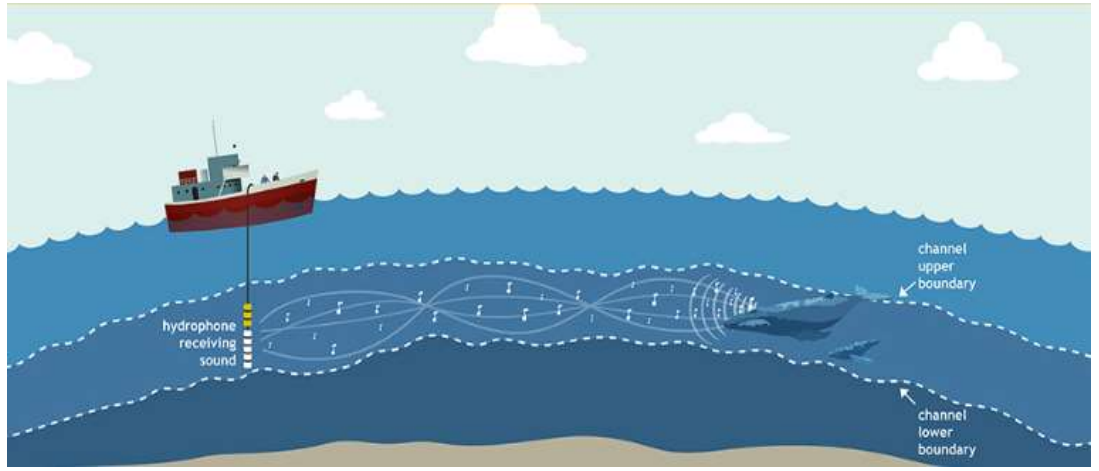
1877 ve 1878'de İngiliz bilim adamı John William Strutt, üçüncü Baron Rayleigh, modern akustik biliminin çalışmalarının başlangıcı olarak kabul edilen iki ciltlik çığır açıcı eseri The Theory of Sound'u yayınladı. Argon elementini başarılı bir şekilde izole etmesiyle 1904'te Nobel Fizik Ödülü'nü alan Lord Rayleigh, akışkanlardaki dalga yayılımı teorisi için kritik olan akustik ve optik alanlarında önemli keşiflerde bulundu. Lord Rayleigh bir ses dalgasını matematiksel bir denklem olarak tanımlayan ilk kişidir (akustik üzerine tüm teorik çalışmaların temeli). 1912'de yaşanan Titanik faciası buzdağları ve diğer seyir tehlikeleri konusunda uyaracak daha aktif cihazlar geliştirmeye araştırmacıları motive etti. İlk işlevsel mesafe ölçer (echo ranger) patenti 1914 yılında

Reginald A. Fessenden tarafından alınmıştır. Fessenden'in yaptığı cihaz düşük frekanslı bir ses sinyali yayan ve sonrasında yankıları dinlemek için alıcı moduna geçen elektrikli bir osilatördür. Bu cihaz su altındaki bir buzdağını 2 mil öteden tespit edebilmiştir. Fakat buzdağının yönünü tam olarak belirleyememiştir. Lord Rayleigh ve daha öncesinde Hollandalı bir astronom olan Willebrord Snell'in çalışmalarına dayanarak, Lichte 1919'da, ışığın bir ortamdan diğerine geçerken kırılması gibi ses dalgalarının da sıcaklık, tuzluluk ve basınçta ufak değişikliklerle karşılaştıklarında kırılacağını kuramlaştırmıştır. Ayrıca okyanus akıntılarının ve mevsim değişikliklerinin ses yayılımını etkileyeceğini öne sürmüştür. MIT (Massachusetts Institute of Technology) ve WHOI (Woods Hole Oceanographic Institution) çalışanı olan Athelstan Spilhaus 1937 yılında batitermograf (BT) adı verilen yeni bir cihaz icat etmiştir. BT sıcaklık sensörü su basıncındaki değişimleri tespit etmeye yönelik bir elemanı bünyesinde barındıran küçük bir torpido biçiminde olan cihazın adıdır. Farklı özelliklere sahip katmanlar arasından geçen ses dalgalarının kırılmaya uğradığı, sesin daha hızlı hareket ettiği bölgeden uzaklaşıp yavaşladığı veya bölgeye doğru büküldüğü keşfedildi. Bu bükülme, daha sıcak ve daha soğuk su katmanları arasındaki ayrım çizgisinin hemen altında konumlanan herhangi bir denizaltının sonar sinyallerine karşı görünmez olmasını sağlayan akustik bir "gölge bölgesi" oluşturur. BT kullanarak akustik gölge bölgelerde bir denizaltıyı saklamak ve denizaltının bu gölge bölgelerde saklanacağını öngörüp beklenen ses kırılmalarını hesaba katarak geminin sonar doğrultusunu buna göre ayarlamak DSH için kritik önem taşımaktadır (Kaharl 1999).



Şekil 2.2 Sıcaklık değişimlerine bağlı sesin kırılma ve bükülmesinden dolayı oluşan gölge bölge (Kaharl 1999)

1943 yılında Columbia Üniversitesi'nden Maurice Ewing ve J. L. Worzel, Ewing'in teorisini test etmek için Bahamalar'da bir deney yaptılar. Bu deneyin sonuçlarını analiz ederken SOFAR (Sound Fixing and Ranging) adını verdikleri bir çeşit ses boru kanalı keşfettiler. Derin ses kanalı olarak da bilinen bu boru hattı, Japon Denizi'ndeki su altı patlamalarından alınan sinyalleri analiz eden Lebedev Fizik Enstitüsü'nden Rus akustikçi Leonid Brekhovskikh tarafından bağımsız olarak keşfedildi (Kaharlı 1999). SOFAR, sesin çok uzak mesafelere taşınmasını sağlayan doğal bir okyanus kanalıdır (National Oceanic and Atmospheric Administration 2024a).



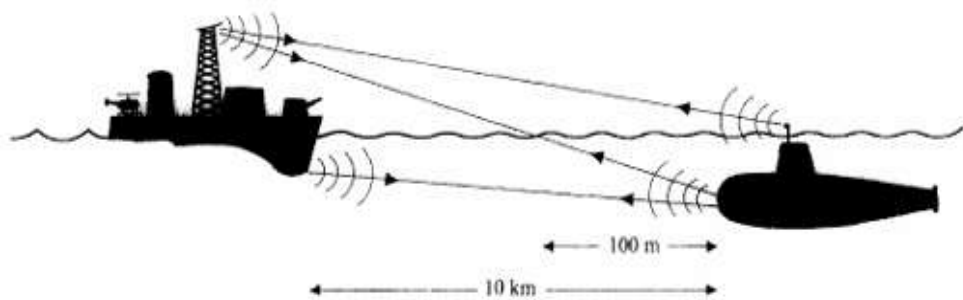
Şekil 2.3 SOFAR kanalı (National Oceanic and Atmospheric Administration 2024a)

1950'lerde gizlilik içinde Amerika Birleşik Devletleri (ABD) Donanması Jezebel kod adıyla bilinen bir proje başlattı. Bu proje daha sonrasında SOSUS (Sound Surveillance System) olarak bilinecekti. SOSUS okyanus tabanına yerleştirilen ve kablolarla kıyıdaki işlem merkezlerine bağlanan hidrofon adı verilen su altı mikrofön dizilerini içermekteydi. Kuzey Amerika'nın her iki kıyısı ile Britanya Batı Hint Adaları boyunca derin ve sığ sularda konuşlandırılan SOSUS sayesinde ABD Donanması, yalnızca Kuzey Yarımküre'nin büyük bölümünde denizaltıları tespit etmekle kalmıyordu; aynı zamanda bir denizaltının kaç pervaneye sahip olduğunu, konvansiyonel mi nükleer mi olduğunu ve bazen hangi sınıfa ait olduğunu dahi ayırt edebiliyordu. Soğuk Savaş'ın sona ermesinin ardından SOSUS'un sivil kullanıma açılması su altı biliminin gelişimine katkı sağlamıştır.

Bir dizi transdüser kullanarak yelpaze biçiminde çok sayıda dar ses dalgası yayan gelişmiş aktif sonar sistemi olan çok ışınlı iskandil (multibeam echosounder), ilk olarak 1960'larda ABD Donanması tarafından geliştirildi. O zamandan beri deniz tabanının yüksek çözünürlüklü ölçümleri için standart teknoloji haline geldi. Deniz tabanının derinliğini tek bir ping ile tek ölçümde belirleyen tek ışınlı iskandilin aksine çok ışınlı iskandil, hüzmleme adı verilen bir sinyal işleme tekniği kullanarak tek bir ping döngüsü içinde yüzlerce batimetrik ölçüm yapabilir. Çok ışınlı sistemler yüksek çözünürlüklü topoğrafya oluşturma'nın yanı sıra deniz tabanının ve su sütununun akustik yansıma ölçümlerini de toplar ve bu da deniz tabanının bileşimi hakkında fikir verir. Bunlara ek olarak su sütunundaki gaz sızıntıları veya balık sürüleri gibi hedeflerin yerini tespit edebilirler (Anonymous 2025b).

1490 yılında Leonardo da Vinci'nin gemilerin su altındaki seslerini dinlemek için boru kullanma önerisinden bu yana, sesin suyun içinde nasıl yayıldığını anlamaya yönelik araştırmalar, sonar teknolojisinin gelişimini şekillendirmiştir (Kaharl 1999). Newton'un ses yayılım teorilerinden Sturm ve Colladon'un suda ses hızını ölçmesine, Lord Rayleigh'nin akustik dalga denklemlerinden Titanik faciasının ardından ivme kazanan mesafe ölçer çalışmalarına, SOFAR kanalı ve SOSUS ağlarından çok ışınlı iskandillere kadar uzanan bu süreç, DSH, iklim değişikliği izleme ve deniz yaşamı araştırmaları gibi alanların gelişimini hızlandırmıştır. Günümüzde, evrimsel sinir ağları (CNN) gibi yapay zekâ yöntemleriyle sonar hedeflerinin otomatik analizi, insan kaynaklı hataları azaltarak bu mirası daha da ileri taşıyacaktır.

## 2.2 Su Altı Hedeflerinin Tespiti için Neden Radar Kullanılmaz?



Şekil 2.4 Su altında neden radar kullanılmaz? (Waite 2002)

Şekil 2.4'te gösterildiği üzere radar (Radio Detection and Ranging) denizaltının periskopunu tespit edecek ve elektromanyetik enerjinin küresel olarak yayıldığını varsayarak ters kare yasası uygulandığında elektromanyetik yayılma kaybı  $40\log 10000 = 160$  dB olacaktır. Sonar ise denizaltının gövdesini tespit edecek ve yine küresel yayılma varsayıldığında sesin yayılım kaybı da 160 dB olacaktır. Bu nedenle su üstünde ve altında radar ve sonar sistemlerinin karşılaştırılabilir performansları vardır. Ancak denizaltı tamamen su altındaysa, son 100 metredeki elektromanyetik kayıplar, herhangi bir radar frekansında iletimi tamamen engelleyecek düzeydedir (Waite 2002).

Yüksek iletkenliğe sahip deniz suyundaki elektromanyetik kayıplar  $1400f^{1/2}$  dB/km ile ifade edilir (f kilohertz cinsindedir). Denizdeki sadece 100 metrelik bir yol için 2000 MHz'deki kayıplar 200000 dB'dir. 30 kHz'de bile kayıplar 770 dB'ye eşittir. Deniz elektromanyetik enerjiye etkili bir şekilde kısa devre oluşturur (Waite 2002).

### 2.3 Zaman-Frekans Ayrıştırma Yöntemleri

Bir ses sinyali genellikle zaman bölgesinde çizilir. Bu şekilde sinyallerin genliğinin zaman içerisinde nasıl değiştiğini incelemek kolaydır. Ancak birçok sinyal değişen frekans özelliklerine sahiptir, örneğin konuşma sinyalleri. Bu özelliği yakalamak için frekans alanında çizilen bir sinyali incelemek faydalıdır. Bu, sinyalin her verilen frekansta ne kadarının bulunduğunu gösteren bir gösterimdir. Sinyalin hem zaman bölgesinde hem de frekans bölgesinde aynı anda incelendiği yöntem olan zaman-frekans ayrıştırma yöntemi, sinyal işlemede yaygın olarak kullanılır. Frekans bölgesinde bir sinyalin gösterimini bulmak için genellikle Fourier dönüşümü kullanılır. Fourier dönüşümünün en yaygın zaman-frekans uygulamalarından biri kısa zamanlı Fourier dönüşümüdür (STFT-Short-Time Fourier Transform) fakat wavelet gibi daha karmaşık yöntemler de yaygın olarak kullanılır.

### **2.3.1 Kısa zamanlı Fourier dönüşümü (Short-time Fourier transform)**

Sinyaller sinüzoidal sinyallerin toplamı olarak gösterilebilir. Temel sinüzoidal sinyalleri belirleyerek, istenenleri tutmak veya çıkarmak mümkündür. Örneğin bir gürültü kaynağı sinüzoidal biçimiyle biliniyorsa izole edilebilir ve çıkarılabilir. STFT, sinyal parametreleri zamanla değişirken bu tür sinüzoidal sinyalleri tespit etmek için kullanılan bir yöntemdir. Bu, bir sinyalin kısa zaman dilimlerine bölünmesi ve her biri üzerinde Fourier dönüşümü uygulanmasıyla yapılır. Bu şekilde her dilim için katsayılar elde edilir. Bu katsayılar, bu kısa zaman dilimlerinin her birinin frekans bölgesi olarak kullanılır. Bu gösterimleri zaman alanında tekrar bir araya getirerek sinyali hem zaman hem de frekans açısından incelemek mümkündür.

### **2.3.2 Dalgacık dönüşümü (Wavelet transform)**

Dalgacık dönüşümünün STFT'den temel farkı, sabit boyutlu bir pencere yerine ölçeklenip ötelenen dalgacık adı verilen kısa ve yerel dalgaları kullanarak bir sinyali temsil etmesidir. Klasik Fourier dönüşümü her frekans için bir katsayı döndürürken, dalgacık dönüşümü her (ölçek, öteleme) çifti için bir katsayı döndürür; böylece sonuç, tek bir frekans parametresi yerine ölçek ve öteleme olmak üzere iki parametrenin fonksiyonu olur. Öteleme, dalgacığın sinyal boyunca zaman eksenindeki konumunu (nereye kaydırılacağını) belirler. Ölçek ise dalgacığın genişleme ve sıkışma miktarını, dolayısıyla analiz ettiği frekans bandını (yalancı frekansını) belirler. Küçük ölçekler yüksek frekanslı detaylara, büyük ölçekler ise düşük frekanslı kaba bileşenlere karşılık gelir.

### **2.3.3 Mel-frekans kepsral katsayıları (Mel-frequency cepstral coefficients)**

Mel-frekans kepsral katsayıları (MFCC-Mel-Frequency Cepstral Coefficients), dönüştürülmüş spektrumda genlik katsayılarını öznelik olarak elde eden gelişmiş bir zaman-frekans analiz tekniğidir. Bu, konuşma ve konuşmacı tanıma uygulamalarında en yaygın kullanılan tekniklerden biridir.

Psikofiziksel çalışmalara göre, insanın seslerin frekans içeriğini algılaması, Mel ölçeği olarak adlandırılan öznel olarak tanımlanmış doğrusal olmayan bir ölçeği takip eder. Mel ölçeği, insan perde algısına dayalı olarak frekansın logaritmik bir ölçeği olarak tanımlanır. Mel birimlerinde eşit aralıklar, eşit perde aralıklarına karşılık gelir. Bu ölçek şu şekilde ifade edilir (Hossan vd. 2010):

$$f_{mel} = 2595 \log_{10} \left( 1 + \frac{f}{700} \right)$$

Burada  $f_{mel}$  Mel birimi cinsinden öznel perdeyi ifade eder ve bu, Hertz (Hz) cinsinden gerçek frekans olan  $f$ ye karşılık gelir (Hossan vd. 2010).

## 2.4 Spektrogram

Spektrogram, belirli bir frekans aralığı için bir sinyalin zaman içindeki gücünü gösteren bir grafikdir. Bir spektrogramın yatay eksen zamanı, dikey eksen ise frekansı gösterir ve her bir frekans-zaman noktasındaki enerji yoğunluğu, renk yoğunluğu veya parlaklıkla ifade edilir. Bu tür görselleştirme, özellikle ses sinyallerinin analizinde yaygın olarak kullanılır ve sesin zamanla nasıl değiştiğini anlamak için güçlü bir araçtır. Spektrogramlar, müzik, sonar, radar, konuşma işleme, sismoloji, elektroensefalogram (EEG) ve elektrokardiyogram (EKG) gibi biyomedikal sinyallerin analizi gibi birçok alanda kullanılmaktadır.

Spektrogramlar, genellikle STFT kullanılarak oluşturulur. Bir  $x(t)$  sinyali için STFT aşağıdaki gibi hesaplanır:

$$STFT(t, f) = \int_{-\infty}^{\infty} [x(\tau) \omega * (\tau - t)] e^{-j2\pi f\tau} d\tau$$

burada  $\omega(t)$ , aşağıdaki koşulu sağlayan pencere fonksiyonunu (window function) ifade etmektedir:

$$||\omega(t)|| = 1$$

STFT'nin ayrık (discrete) formu ise aşağıdaki gibi hesaplanabilir:

$$STFT(n, k) = \sum_{m=-\infty}^{\infty} [x(m)\omega * (m - n)]e^{-j2\pi km}$$

burada  $\omega(t)$ , ayrık alandaki (discrete domain) zaman penceresi fonksiyonunu belirtmektedir (Ghavidel vd. 2022).

### **3. SONAR TEKNOLOJİSİNİN UYGULAMA ALANLARI VE DENİZCİLİK SEKTÖRÜNDEKİ KONUMU**

Bu bölümde sonar teknolojisinin küresel çaptaki kullanım alanları, denizcilik sektöründeki özgül uygulamaları ve Türkiye'nin bu alandaki mevcut konumu incelenmektedir.

#### **3.1 Sonar Teknolojisinin Uygulama Alanları**

Sonar teknolojisi askeri ve sivil olmak üzere iki alanda incelenebilir.

##### **3.1.1 Askeri uygulamalar**

Sonar, deniz kuvvetleri için vazgeçilmez bir stratejik araç olup su altı hedeflerini tespit etme, takip etme ve sınıflandırmada kritik bir rol oynamaktadır.

Askeri uygulamaların detaylı incelenmesi, modern DSH'nin tek bir sonar sistemine dayanmadığını, aksine çok çeşitli sonar türlerini (karinaya monteli, çekili dizin, daldırma, değişken derinlik) kapsadığını göstermektedir. Her bir sonar türü, farklı operasyonel senaryolar ve platformlar için kendine özgü avantajlar sağlamaktadır. Ayrıca, bu sistemler sadece kendi aralarında değil, aynı zamanda aktif karşı tedbirler ve özel eğitim modülleri ile de giderek daha fazla entegre edilmektedir. Bu karmaşık etkileşim, bireysel bileşenlerin daha büyük, daha yetenekli bir bütüne katkıda bulunduğu bir "sistemler sistemi" yaklaşımını işaret etmektedir.

Türkiye için, DSH'de ve genel deniz güvenliğinde operasyonel üstünlük elde etmek ve bu üstünlüğü sürdürmek, sadece bireysel yüksek performanslı sonar birimlerini tedarik etmek veya üretmekten öte bir geliştirme stratejisi gerektirmektedir. Odak noktası, su üstü gemileri, denizaltılar, helikopterler veya insansız deniz araçları üzerinde konuşlandırılmış tüm sonar varlıkları arasında sorunsuz birlikte çalışabilirlik, sağlam veri füzyonu ve gerçek zamanlı bilgi alışverişi sağlamak olmalıdır. Gelişmiş savaş yönetim sistemleri tarafından kolaylaştırılan bu bütünsel yaklaşım, Türk Deniz Kuvvetleri'nin kapsamlı ve

dinamik bir su altı taktik resmi oluřturmasına olanak tanıyacak, durumsal farkındalıęını, tepki yeteneklerini ve Mavi Vatan doktrini kapsamındaki deniz yetki alanlarının korunmasındaki genel etkinlięini önemli ölçüde artıracaktır.

Sonar teknolojisinin başlıca askeri kullanım alanları ařaęıdaki başlıklarda özetlenmiřtir.

#### **3.1.1.1 Denizaltı savunma harbi**

Denizaltıların tespiti, takibi ve sınıflandırılması, aktif ve pasif sonar sistemlerinin en temel askeri uygulamasıdır. Özellikle modern, sessiz dizel-elektrik denizaltılarının oluřturduęu tehdide karřı, düşük frekanslı aktif sonarlar gibi uzun menzilli sistemler geliřtirilmiřtir. Bu sistemler, sıę ve derin sularda tutarlı tespit performansı sergilemektedir.

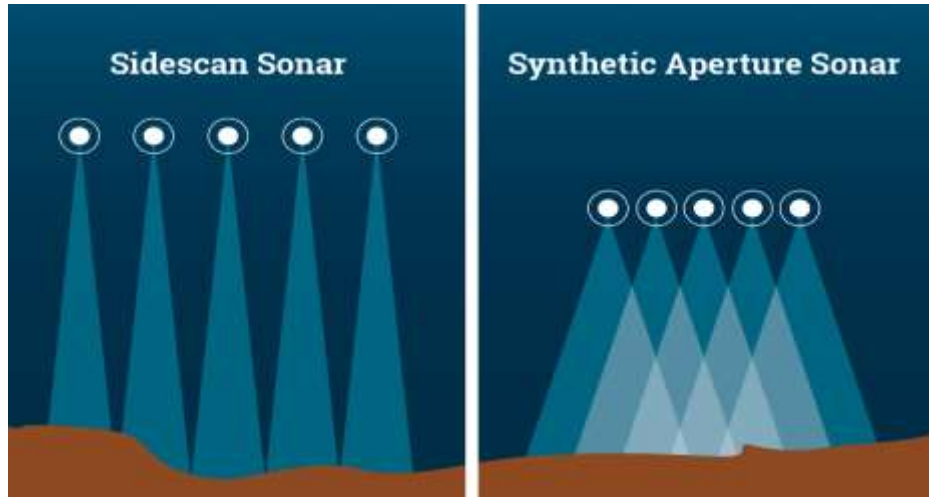
#### **3.1.1.2 Mayın tespiti ve karřı tedbirler**

Deniz tabanındaki mayınların tespiti, sınıflandırılması ve etkisiz hale getirilmesi için yandan taramalı sonarlar (Side-Scan Sonar) ve yüksek çözünürlüklü sentetik açıklıklı sonarlar (SAS-Synthetic Aperture Sonar) kullanılmaktadır. SAS, özellikle mayın avcılıęında ve deniz altı gözetiminde kritik öneme sahip olup yüksek çözünürlüklü görüntüleme yeteneęi sunar.

Geleneksel yandan taramalı sonar sistemlerinde, transdüser dizisi bir akustik darbe gönderir ve bir sonraki darbeyi göndermeden önce yankının geri dönmesini bekler. Buna karřılık sentetik açıklıklı sonar (SAS), geri dönen sinyalleri işlemeksizin sürekli darbeler gönderir ve bir bölgeden alınan çoklu ölçümleri birleřtirerek yapay olarak diziye uzatır. SAS demeti, aynı konumu çoklu kez örnekleyerek üst üste binen bir “huni” etkisi oluřturarak yüksek ayrıntı seviyesinde görüntüler elde eder. Bu sayede sentetik açıklıklı sonar, bir bölgeyi geleneksel yandan taramalı sonardan 30 kata kadar daha yüksek çözünürlükle haritalayabilmektedir (National Oceanic and Atmospheric Administration 2024c).



Şekil 3.1 Nantucket kıyılarında kimliği belirlenemeyen bir gemi enkazının SAS taraması (National Oceanic and Atmospheric Administration 2024c)



Şekil 3.2 Yandan taramalı sonar ve sentetik açıklıklı sonar çalışma yapısının karşılaştırılması (National Oceanic and Atmospheric Administration 2024c)

### 3.1.1.3 Liman ve kritik tesis güvenliği

Limanlara, deniz üslerine ve kritik deniz altyapılarına yönelik dalgıçlar, su altı dronu olarak da bilinen insansız su altı araçları (UUV-Unmanned Underwater Vehicle) gibi su altı tehditlerini tespit etmek amacıyla dalgıç tespit sonarları (DTS) kullanılmaktadır.

#### **3.1.1.4 Su altı gözetleme ve keşif**

Uzun menzilli ve geniş alan gözetlemesi için çekili dizin sonarlar (TAS-Towed Array Sonar) ve değişken derinlikli sonarlar (VDS-Variable Depth Sonar) tercih edilmektedir. Otonom su altı araçları (AUV-Autonomous Underwater Vehicle) üzerine entegre edilen sonarlar da keşif ve gözetleme görevlerinde giderek daha fazla kullanılmaktadır.

#### **3.1.1.5 Torpidoya karşı tedbirler**

Bazı gelişmiş sonar sistemleri, yaklaşan torpido tehditlerini algılama ve otomatik olarak karşı tedbir ekipmanlarını (örneğin akustik karıştırıcılar ve aldaticılar) devreye sokma yeteneğine sahiptir.

#### **3.1.1.6 Denizaltı eğitimi**

Sonar Transponder Modülleri gibi sistemler, denizaltıların akustik imzalarını ve yankı tepkilerini gerçekçi bir şekilde simüle ederek, DSH eğitimlerinde maliyet etkin ve tekrar kullanılabilir çözümler sunmaktadır.

#### **3.1.2 Sivil uygulamalar**

Sonar teknolojisi, askeri alanların ötesinde geniş bir sivil uygulama yelpazesine sahiptir ve günlük yaşamdan bilimsel araştırmalara kadar birçok alanda kritik roller üstlenmektedir.

Sonar teknolojisinin askeri uygulamaları ön planda olsa da balıkçılık, oşinografi, petrol/gaz keşfi, arama ve kurtarma, tıbbi teşhis ve endüstriyel izleme gibi kapsamlı ve çeşitli sivil uygulamalar, bu teknolojinin küresel ekonominin ve insan refahının çeşitli sektörleri üzerinde derin ve yaygın bir etkisi olduğunu göstermektedir. Bu sadece niş kullanımlarla ilgili değil, sürdürülebilir kaynak yönetimi, güvenliğin artırılması, sağlık hizmetlerinin iyileştirilmesi ve gezegenimiz hakkındaki bilimsel anlayışın

derinleştirilmesini sağlayan temel araçlarla ilgilidir. Küresel sonar sistemleri pazarındaki beklenen büyüme, bu çeşitli uygulamalar genelinde artan ekonomik önemini daha da vurgulamaktadır.

Türkiye için, sonar teknolojisine, özellikle yerli geliştirmeye yönelik sürekli yatırım, ülkeyi sadece zorlu bir savunma gücü olarak değil, aynı zamanda çeşitli yüksek değerli ticari ve bilimsel sektörlerde potansiyel bir lider olarak konumlandırmaktadır. Bu geniş uygulanabilirlik, ulusal Ar-Ge çabaları için yatırım getirisini önemli ölçüde artırmakta, daha sağlam ve çeşitlendirilmiş bir teknolojik taban oluşturmakta ve kritik sivil altyapı ve endüstriler için dış tedarikçilere olan bağımlılığı azaltmaktadır. Bu çok sektörlü fayda, genel ulusal inovasyon ekosistemini güçlendirmekte ve daha dirençli bir ekonomiye katkıda bulunmaktadır.

Sonar teknolojisinin başlıca sivil kullanım alanları aşağıdaki başlıklarda özetlenmiştir.

### **3.1.2.1 Balıkçılık ve su ürünleri yetiştiriciliği**

Sivil alanda en yaygın kullanılan sonarlar balıkçılık alanında kullanılan sonarlardır. Sonar teknolojisinin balıkçılıkta kullanımı, balıkçılara balıkların ve su altı yapıların yerini hassas biçimde belirleme imkânı sunarak av başarılarını artırmaktadır. Balıkçılıkta kullanılabilecek çeşitli sonar teknolojileri bulunmaktadır. Bunlar arasında, balıkların varlığını algılayarak ekranda gösteren balık bulucu sonarlar, su altı yapısını ayrıntılı şekilde görüntüleyen yandan taramalı sonar sistemleri ile taşınabilir ve elde kullanılabilen balık bulucular yer almakta olup bu sayede kullanım kolaylığı ve esneklik sağlanmaktadır.

### **3.1.2.2 Oşinografik araştırmalar ve deniz tabanı haritalaması**

Deniz derinliğinin hassas bir şekilde ölçülmesi, deniz tabanı topografyasının detaylı haritalarının oluşturulması, sediman ve deniz yatağı tiplerinin belirlenmesi için tek hüzmeli (single beam) ve çok hüzmeli iskandiller (multibeam echosounders) ile yandan

taramalı sonarlar kullanılır. Bu uygulamalar, okyanus akıntılarının izlenmesi, deniz ekosistemlerinin anlaşılması ve iklim değişikliğinin okyanuslar üzerindeki etkilerinin araştırılması için hayati öneme sahiptir.

#### **3.1.2.3 Petrol ve gaz keşfi**

Deniz tabanının altındaki jeolojik yapıların, potansiyel petrol ve gaz yataklarının tespiti için sismik yansıma ve kırılma teknikleri kullanılır. Bu süreçte, basınçlı hava tabancaları ses kaynağı olarak kullanılır ve çekili hidrofon dizileri yansıyan sinyalleri toplar.

#### **3.1.2.4 Arama ve kurtarma operasyonları**

Su altında kayıp nesnelerin, batık araçların veya kişilerin bulunması için sonar sistemleri (özellikle yandan taramalı sonarlar) ve uzaktan kumandalı su altı araçlarına (ROV- Remotely Operated Vehicle) entegre sonarlar kullanılır. Bulanık ve düşük görüş mesafesindeki su ortamında dahi detaylı görüntüler sağlayarak arama ve kurtarma operasyonlarının etkinliğini ve güvenliğini artırır.

#### **3.1.2.5 Su altı altyapı denetimi ve bakımı**

Köprüler, barajlar, boru hatları ve diğer deniz altı yapıları gibi kritik altyapıların hasar tespiti, yapısal sorunlarının incelenmesi ve düzenli bakımı için ROV'lara entegre sonarlar kullanılmaktadır.

#### **3.1.2.6 Tıbbi uygulamalar**

Ultrasonografi teknolojisi, insan vücudunun iç yapılarının (organlar, kan damarları, fetüs gelişimi) radyasyon içermeyen görüntülenmesi ve tanısal amaçlar için yaygın olarak kullanılmaktadır.

### **3.1.2.7 Endüstriyel izleme ve tahmini bakım**

Endüstriyel tesislerde makinelerin ve ekipmanların durumunu izlemek, potansiyel arızaları tahmin etmek ve önleyici bakım yapmak için akustik sensörler kullanılır.

## **3.2 Denizcilik Sektöründe Kullanılan Sonar Konfigürasyonları**

Denizcilik sektöründe farklı operasyonel ihtiyaçlara ve çevresel koşullara göre çeşitli sonar konfigürasyonları kullanılmaktadır.

Gerçek su altı alan farkındalığına ulaşmak, sonar sistemlerinin farklı platformlar (su üstü, su altı, hava, insansız) arasında entegrasyonunu gerektirmektedir. Bu kapsamlı entegrasyon, katmanlı bir savunma, daha iyi hedef lokalizasyonu ve artırılmış durumsal farkındalık sağlar. Farklı sonar konfigürasyonlarının bir araya gelmesi, deniz kuvvetlerinin geniş bir yelpazedeki su altı tehditlerine karşı daha etkili bir şekilde mücadele etmesini mümkün kılmaktadır.

Denizcilik sektöründe kullanılan sonar konfigürasyonlarından başlıcaları aşağıdaki başlıklarda özetlenmiştir.

### **3.2.1 Karinaya monteli sonar**

Geminin gövdesine veya pruvasına doğrudan sabitlenen sonarlardır. Genel su altı tespiti, tehlike izleme ve torpido saldırısı uyarıları için geminin ilk savunma hattı olarak işlev görürler. Derinlik kısıtlamaları olsa da çoklu konfigürasyonlarda çalışabilmeleri nedeniyle çoğu savaş gemisinde standart olarak bulunur.

### **3.2.2 Çekili dizin sonar (TAS-Towed Array Sonar)**

Geminin arkasından, bir kablo ile çekilen uzun hidrofon dizinleridir. Bu konfigürasyon, dizinin geminin kendi gürültü kaynaklarından uzakta tutulmasını sağlayarak sinyal-

gürültü oranını (SNR-Signal-to-Noise Ratio) önemli ölçüde iyileştirir ve sessiz denizaltı tehditleri gibi zayıf temasların uzun menzilde tespit ve takibini mümkün kılar. Ayrıca, termal katmanların altındaki hedefleri tespit etmek için dizinin derinliği ayarlanabilir. Üstün menzil ve çözünürlük sunmasına rağmen, sistemin etkin kullanımı geminin hızını sınırlar ve kablonun hasar görme riski bulunmaktadır.

### **3.2.3 Değişken derinlikli sonar (VDS-Variable Depth Sonar)**

Termal katmanlardaki denizaltı takibi için çeşitli derinliklere indirilebilen sistemlerdir. Bu sistemler, sesin farklı su katmanlarında farklı yayılım özelliklerine sahip olmasından kaynaklanan dezavantajları telafi ederek, denizaltıların termal katmanların altında gizlenmesini zorlaştırır. Özellikle kıyı sularında ve güçlü termal gradyanların olduğu bölgelerde denizaltı avcılığı için kritik bir araçtır.

### **3.2.4 Daldırma sonarı**

Helikopterlerden suya indirilen ve şüpheli alanların hızlı taranması için kullanılan sonarlardır. Bu sonar tipi dalgıçlar, insansız su altı araçları ve küçük denizaltılar gibi küçük ve tespit edilmesi güç hedefleri tespit etme yeteneğine sahiptir.

### **3.2.5 Yandan taramalı sonar**

Deniz tabanını yanal olarak ses hüzmeleriyle tarayarak deniz tabanının şeklini ve dokusunu ortaya çıkaran detaylı görüntüler oluşturur. Deniz tabanı görüntülemesi, arama ve kurtarma, deniz inşaatı ve arkeoloji için kullanılır. Geniş alanları hızlı ve verimli bir şekilde kapsayabilir.

### **3.2.6 Sentetik açıklıklı sonar (SAS-Synthetic Aperture Sonar)**

Yapay bir dizin kullanarak yüksek çözünürlüklü görüntüler yakalayan gelişmekte olan bir sonar türüdür. Geleneksel yandan taramalı sonara göre 30 kata kadar daha yüksek

özünörlökte haritalama yapabilir. Mayın avcılığı, denizaltı gözetimi, költörel miras alanlarının (batıklar) görüntölenmesi ve deniz tabanı sediman yapısının karakterizasyonu için önemlidir. Genellikle otonom su altı araçlarıyla (AUV) eşleştirilir ve hassas haritalama ve su altı nesnelerinin tespiti için kullanılır. SAS'ın yüksek özünörlö yeteneđi, her bir alanın birden fazla kez ölçölmesi ve sinyallerin birleştirilmesiyle yapay olarak dizinin açıklılıının uzatılması prensibine dayanır. Dezavantajları arasında yüksek işletme maliyeti, yandan taramalı sonardan genellikle daha az alan kapsaması ve hassas bir hareket yöröngesi gerektirmesi yer alır.

### **3.2.7 İskandil (Echo Sounder)**

Deniz tabanının derinliliğini ölçmek için kullanılan sonar sistemleridir. Ses dalgalarının deniz tabanına gönderilip geri dönmesiyle geçen süreyi ölçerek derinlik hesaplaması yapar. Navigasyon, haritalama ve deniz tabanı topografyasının belirlenmesi için temel bir araçtır. Tek ışınlı (single-beam) ve ok ışınlı (multi-beam) iskandil sistemleri olmak üzere iki ana türü bulunur. Tek ışınlı sistemler, daha basit ve ekonomik olup genellikle küçük alanların taranmasında kullanılırken, ok ışınlı sistemler geniş alanlarda yüksek özünörlöklö derinlik haritaları oluşturmak için tercih edilir. İskandil sistemleri, balıkçılık, hidrografik araştırmalar ve deniz tabanı haritalandırması gibi sivil uygulamaların yanı sıra askeri amaçlarla da kullanılır.

## 4. MATERYAL VE YÖNTEM

Bu bölüm, su altı akustik hedef tanıma problemine yönelik olarak geliştirilen CNN modelinin materyal ve yöntemlerini açıklamaktadır. Çalışmanın tekrarlanabilirliğini ve güvenilirliğini sağlamak amacıyla, kullanılan veri seti, ön işleme adımları, model mimarisi, eğitim kriterleri ve performans değerlendirme metrikleri açıklanmıştır.

### 4.1 Giriş ve Problem Tanımı

Su altı akustik hedef sınıflandırma hem askeri hem de sivil uygulamalarda kritik bir rol oynamaktadır. Askeri alanda, dost ve düşman hedeflerinin belirlenmesi ve DSH gibi stratejik görevler için vazgeçilmezdir. Sivil uygulamalarda ise çevresel izleme, balıkçılık ve deniz kirliliğinin tespiti gibi alanlarda önemli katkılar sağlamaktadır.

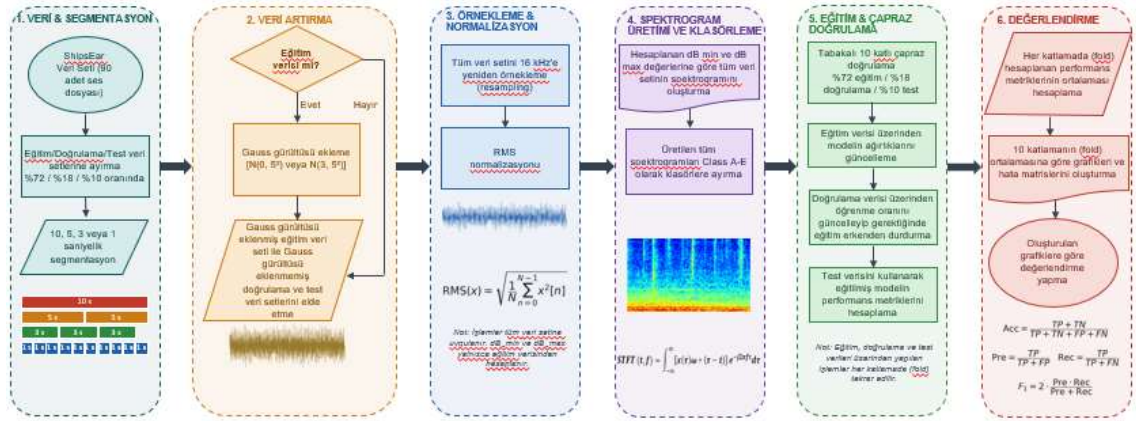
Su altı akustik hedef sınıflandırma alanında karşılaşılan temel zorluklar, bu teknolojinin geliştirilmesini karmaşık hale getirmektedir. Yüksek kaliteli ve etiketli su altı akustik verisi elde etmenin maliyeti oldukça yüksektir ve bu süreç büyük zorluklar içermektedir. Askeri verilerin gizliliği nedeniyle açık kaynak veri setlerinin sınırlı olması, araştırmacıların bu alanda çalışmasını kısıtlamaktadır. Ayrıca, su altı ortamının doğası gereği karmaşık olması, arka plan gürültüsü (doğal, biyolojik ve insan kaynaklı), frekansa bağlı soğurma, çoklu yol etkileri (yansıma, kırılma, saçılma) ve Doppler etkisi gibi faktörler sinyal işlemeyi zorlaştırmaktadır. Gemilerden yayılan gürültünün zamanla değişen, doğrusal olmayan, salınımlı ve durağan olmayan yapısı, sinyal karakteristiklerinin değişkenliğini artırmaktadır. Geleneksel olarak su altı akustik hedef tanıma, sonar operatörlerinin uzmanlığına dayanmaktadır, ancak bu yaklaşım sürekli izleme gerektirmesi, insan faktörlerinden (ruh hali, fiziksel durum) etkilenmesi ve maliyetli eğitim süreçleri nedeniyle sınırlılıklara sahiptir. Bu zorluklar, su altı akustik hedef tanımayı sadece bir sinyal işleme problemi olmaktan çıkarıp, aynı zamanda bir veri mühendisliği, makine öğrenmesi ve çevresel modelleme problemi haline getirmektedir. Veri kıtlığı, derin öğrenme modellerinin genelleme yeteneğini doğrudan etkileyen en büyük engellerden biridir. Sınırlı ve dengesiz dağılımlı veri setleri, modellerin veri

açlığını tetikleyerek aşırı öğrenme riskini artırır ve gerçek dünya uygulamalarında düşük performansa yol açabilir.

Son yıllarda derin öğrenme yaklaşımlarının farklı alanlardaki kayda değer başarıları, su altı akustik hedef tanıma alanına uygulanması için güçlü bir ilham kaynağı olmuştur. Özellikle CNN gibi modeller, su altı akustik sinyallerindeki karmaşık örüntüleri otomatik olarak öğrenme yeteneği sunmaktadır. Geleneksel özellik çıkarım metotlarından (LOFAR, DEMON) derin öğrenmenin otomatik özellik çıkarımına geçiş, karmaşık ve gürültülü su altı verileri için daha sağlam ve uyum sağlayabilir çözümler sunmaktadır. Bu paradigma değişimi, insan uzmanlığına olan bağımlılığı azaltma potansiyeli taşımaktadır. Geleneksel yöntemlerin sınırlılıkları ve gürültüye karşı zayıf direnci derin öğrenmenin karmaşık örüntüleri otomatik öğrenme yeteneği ile aşılabilmektedir.

Bu çalışmanın temel amacı, ShipsEar veri kümesi üzerinde spektrogram tabanlı CNN modellerinin su altı akustik hedef sınıflandırma performansını kapsamlı bir şekilde analiz etmektir. Özellikle, farklı segmentasyon uzunlukları ile eğitim setine uygulanan Gauss gürültüsü artırma rejimlerinin, modelin öğrenme dinamikleri ve genelleme yeteneği üzerindeki etkileri iki faktörlü bir tasarım çerçevesinde karşılaştırmalı olarak gösterilmiştir. Çalışma, EfficientNetB1 gibi önceden eğitilmiş mimarilerin su altı akustik hedef sınıflandırma problemine uygulanabilirliğini ve performansını değerlendirmeyi amaçlamaktadır.

Bu bölümde ayrıntılandırılan yöntemin bütününi oluşturan işlem adımları ve bunların birbiriyle ilişkisi Şekil 4.1’de özetlenmiştir. Süreç, birbirini izleyen altı temel aşamadan meydana gelmektedir: (i) veri hazırlama ve segmentasyon, (ii) veri artırma, (iii) yeniden örnekleme ve normalizasyon, (iv) spektrogram üretimi ve sınıflara ayırma, (v) model eğitimi ile çapraz doğrulama ve (vi) başarımlı değerlendirme. Şekil 4.1, aynı zamanda çalışmanın iki faktörlü tasarımını (segment uzunluğu ve eğitim setine uygulanan Gauss gürültüsü rejimi) görsel olarak belirginleştirmekte; izleyen alt bölümler (Bölüm 4.2–4.6) ise her bir aşamanın ayrıntılarını sunmaktadır.



Şekil 4.1 Çalışmada izlenen metodolojinin genel akış şeması

## 4.2 Veri Seti ve Ön İşleme

Bu çalışmada kullanılan ShipsEar veri seti, 2012-2013 yılları arasında kuzeybatı İspanya'nın Atlantik kıyısı boyunca seyahat eden çeşitli gemilerden elde edilen ses kayıtlarını içermektedir. Kayıtlar, digitalHyd SR-1 hidrofonu kullanılarak alınmıştır. Bu hidrofonun örnekleme frekansı 52734 Hz ve örnekleme çözünürlüğü 24 bittir. 11 farklı gemi tipi ve arka plan gürültülerinden oluşan ShipsEar veri seti, en kısa kaydı 9 saniye, en uzununu ise 11 dakika olmak üzere toplamı 3 saat 9 dakika olan 90 adet ses kaydı içermektedir (Santos-Domínguez vd. 2016). Bahsedilen veri seti, bu çalışmada Class A, Class B, Class C, Class D ve Class E olmak üzere 5 ana sınıfa ayrılmıştır. Gemi büyüklüğüne göre yapılan bu sınıflara ayırma işlemi Çizelge 4.1'de verilmiştir.

Çizelge 4.1 Oluşturulan 5 sınıfın ayrıntılı açıklaması

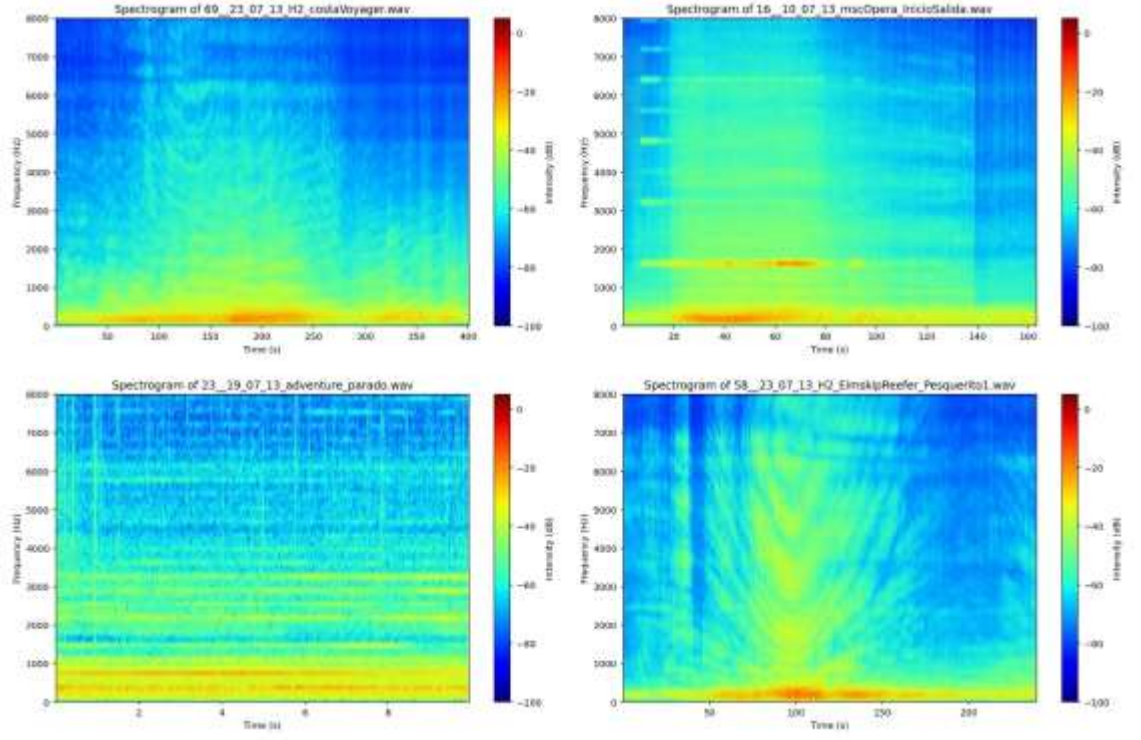
| Sınıf   | Tip                                                                  |
|---------|----------------------------------------------------------------------|
| Class A | Okyanus gemisi, RORO                                                 |
| Class B | Balıkçı teknesi, Trol teknesi, Midye teknesi, Römorkör, Tarak gemisi |
| Class C | Motorlu tekne, Kılavuz botu, Yelkenli tekne                          |
| Class D | Yolcu gemisi                                                         |
| Class E | Arka plan gürültüsü                                                  |

Ses dosyalarındaki segmentasyon yöntemi uygulanmadan önce 90 adet ses kaydının spektrogramları çıkarılmış olup 5 farklı sınıfa ait spektrogramların oluşturulma yöntemi açıklanmış ve sınıfsal bazda incelenmiştir.

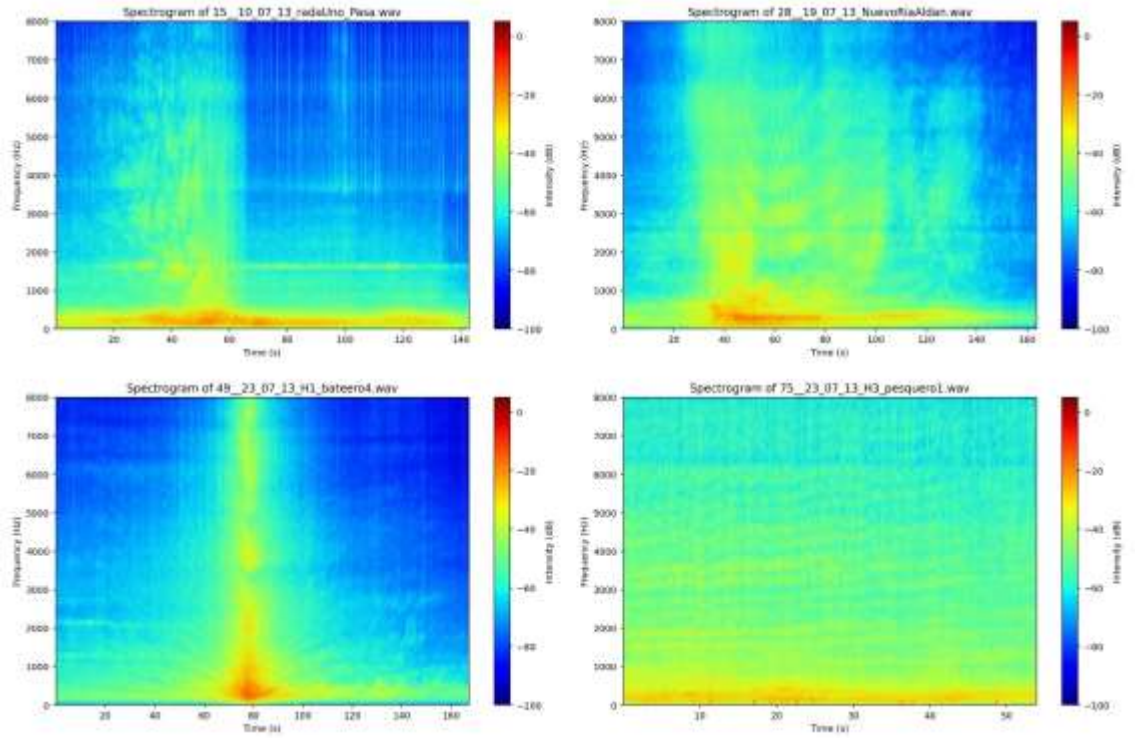
Veri setinin sınırlı olması nedeniyle, Xu ve diğerleri tarafından raporlanan veri çoğaltma yaklaşımı uyarlanarak ham ses sinyallerine Gauss gürültüsü eklenmiştir (Xu vd. 2024). Çoğaltma işlemi yalnızca eğitim setine uygulanmış olup doğrulama ve test setleri ham halleriyle korunmuştur. Bu sayede eğitim seti dört kat artırılmıştır. Bu ayrımın amacı, test doğruluğunun modelin sınıflandırma kabiliyetini yansıtmasını sağlamak ve eklenen gürültü dağılımına aşırı uyumun oluşmasını engellemektir. Çalışmada gürültü artırımı, tek bir sabit ayar yerine iki ayrı rejim olarak ele alınmıştır: 0 dB ortalama ve 5 dB standart sapmaya sahip  $N(0, 5^2)$  rejimi ile 3 dB ortalama ve 5 dB standart sapmaya sahip  $N(3, 5^2)$  rejimi. Ayrıca gürültü artırımının başarıma net katkısını yalıtılabilmek amacıyla, 5 saniyelik segmentasyon için gürültünün hiç eklenmediği bir kontrol deneyi de yürütülmüştür. Gürültü rejiminin sınıflandırma başarımı üzerindeki etkisi Bölüm 4.6’da ayrıntılı olarak değerlendirilmektedir.

Öncelikle bu çalışmada ShipsEar veri setinin özgün örnekleme frekansı olan 52734 Hz, hesaplama verimliliğini artırmak amacıyla 16 kHz’e düşürülmüştür (Yang vd. 2024). Nyquist örnekleme teoremi gereği bu işlem sonucunda analiz edilebilir maksimum frekans 8 kHz ile sınırlanmaktadır. Su altı akustik hedef tanımda daha ayrıntılı öznelik bilgisinin genellikle düşük frekanslarda elde edildiği bilinmektedir (Yang vd. 2024). En iyi sınıflandırma oranı için ilk 8 kHz’lik kısım akustik bilgi açısından en anlamlı frekans bandı olup 8 kHz üzerindeki frekans bileşenlerinin analiz edilmesi deneysel anlamda herhangi bir iyileştirme sağlamamaktadır (Santos-Domínguez vd. 2016). İkinci olarak bu çalışmada, farklı kayıt koşulları ve mesafelerden kaynaklanan sinyal enerjisi farklılıklarını gidermek amacıyla RMS normalizasyonu uygulanmıştır (Liu vd. 2021).

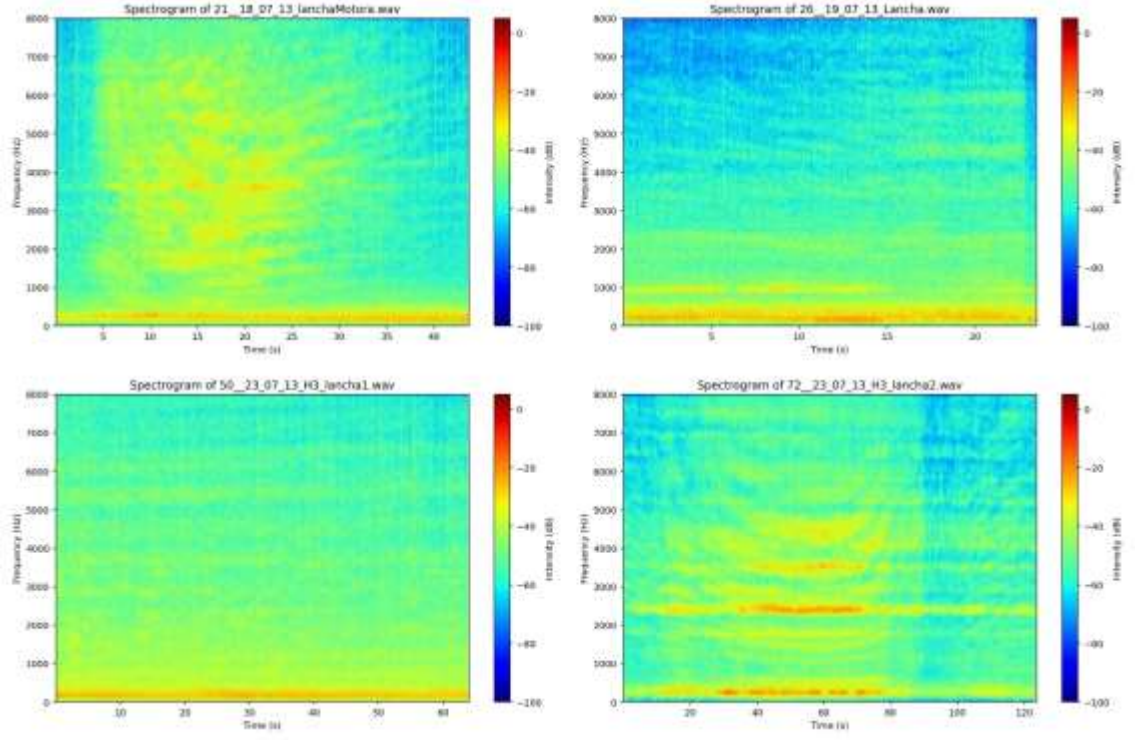
Son olarak bu çalışmada, eğitim seti taranarak belirlenen dB cinsinden sinyalin global minimum ve maksimum güç değerleri sabit tutularak STFT aracılığıyla spektrogramlar elde edilmiştir. Bu yaklaşımla tüm spektrogramların aynı renk skalasında görselleştirilmesi ve sınıflar arası karşılaştırılabilirlik sağlanmıştır.



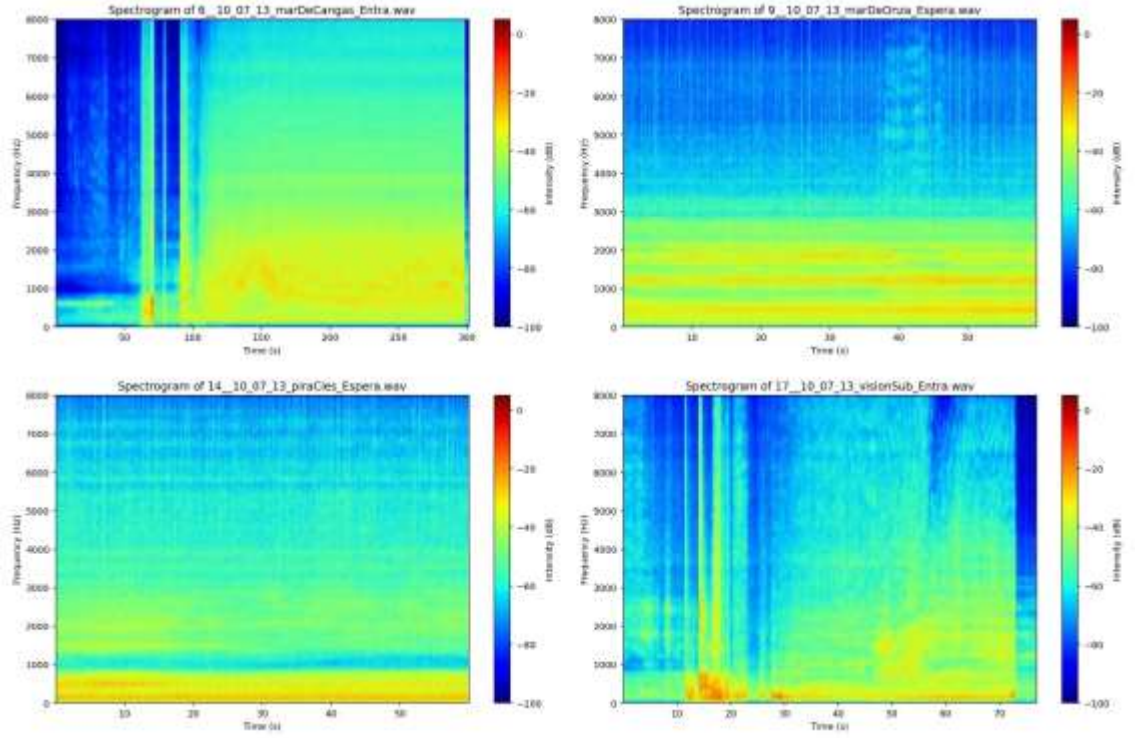
Şekil 4.2 Class A'ya ait örnek spektrogramlar



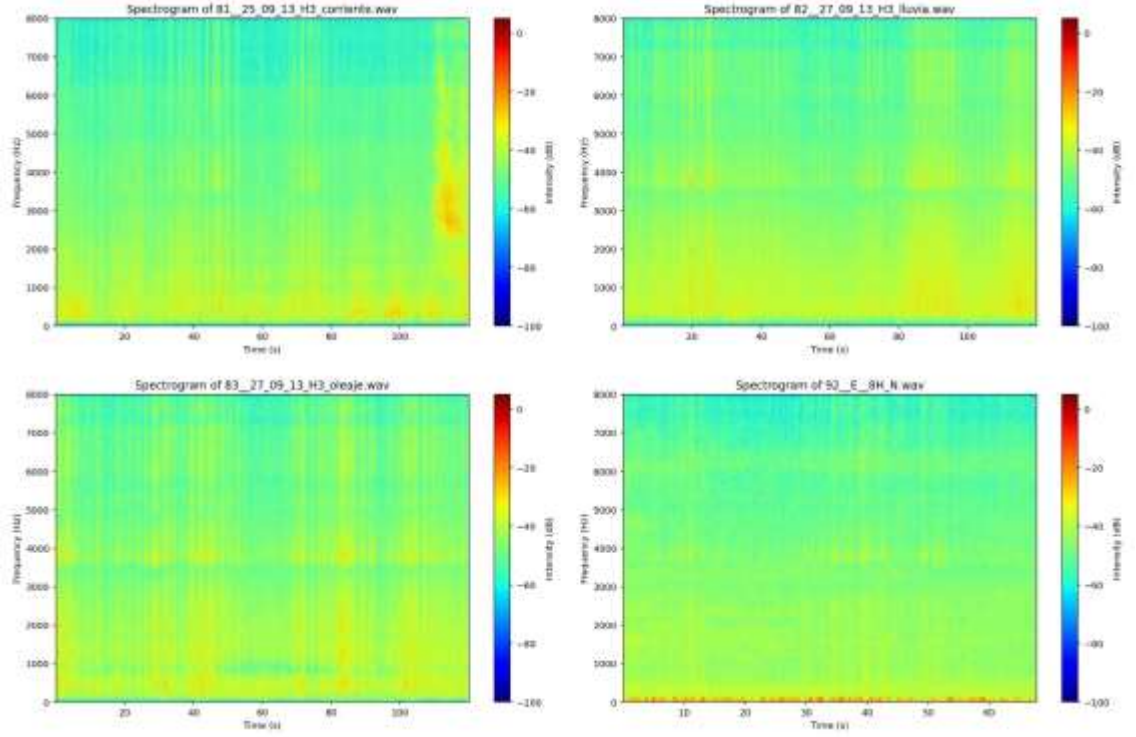
Şekil 4.3 Class B'ye ait örnek spektrogramlar



Şekil 4.4 Class C'ye ait örnek spektrogramlar



Şekil 4.5 Class D'ye ait örnek spektrogramlar



Şekil 4.6 Class E'ye ait örnek spektrogramlar

ShipsEar veri setine ait beş sınıfın spektrogramları incelendiğinde, sınıflar arasındaki akustik farklılıkların spektral yapı, tonal bileşen dağılımı ve zamansal değişkenlik açısından belirgin biçimde kendini gösterdiği görülmektedir. Şekil 4.2'de Class A sınıfındaki büyük tonajlı gemilere (okyanus gemisi, RORO) ait spektrogramlarda düşük frekans bandı boyunca yüksek enerji yoğunluğu ve belirgin tonal hat yapıları gözlemlenmiştir. Zhu ve diğerleri (Zhu vd. 2022), bir araştırma gemisine ait spektrogram verilerini analiz ederek gemi gürültüsünün 0–3 kHz arasında baskın olduğunu ve gemi tarafından yayılan sinyalin dar bantlı makine tonal bileşenleri ile geniş bantlı pervane kavitasyon gürültüsünü eş zamanlı olarak barındırdığını göstermiştir. Class B'de yer alan balıkçı teknesi, trol ve römorkör gibi orta ölçekli deniz araçlarına ait spektrogramlarda ise ani enerji başlangıçları ve belirgin zamansal değişkenlikler dikkat çekmektedir. Shipton ve diğerleri (Shipton vd. 2025), gemilerin su altına yaydığı akustik gürültü profilinin; motor ateşlemesi sırasında meydana gelen yüksek enerjili anlık darbe dalgalarını, itici sistemler, ana makineler, pompalar ve jeneratörler tarafından üretilen dar bantlı gürültüleri ve kavitasyon gürültülerini kapsadığını belirtmiştir. Class C'ye ait küçük motorlu teknelerin (motorlu tekne, kılavuz botu, yelkenli tekne) spektrogramlarında 0–4

kHz bandında yoğunlaşan enerji dağılımı ve bu bant içinde belirgin harmonik hat yapıları gözlemlenmiş olup bu durum 25 metreden kısa teknelerin büyük çoğunluğunun, hızlarına bağlı biçimde 1–10 kHz arasında geniş bantlı yüksek frekanslı gürültü ürettiği bulgusuyla örtüşmektedir (Anonymous 2025a). Class D kapsamındaki yolcu gemilerine ait spektrogramlar, bekleme konumunda görece düşük enerji, seyir halinde ise düşük-orta frekans bandında enerji yoğunluğu sergilemiştir. Bu örüntü, sabit güzergahlarda çalışan gemilerin tahrik sistemlerinin kararlı çalışma rejimini yansıtmaktadır. Class E'yi oluşturan doğal ortam gürültüsü kayıtları ise diğer dört sınıftan tonal hat veya sistematik harmonik yapı içermemesi bakımından ayrışmakta; tüm analiz bandında homojen dağılmış geniş bantlı stokastik bir zemin sergilemektedir. Sınıflar arasında gözlemlenen bu farklılıklar, ShipsEar veri setinin CNN tabanlı sınıflandırma için yeterli akustik ayrıştırılabilirliğe sahip olduğuna işaret etmekte ve görsel temsili olan spektrogramların girdi olarak kullanımını desteklemektedir (Zhu vd. 2022).

Parlaklık, kontrast, doygunluk ve renk tonu değişimlerini kapsayan color jitter gibi fotometrik veri artırımı (photometric data augmentation) teknikleri bu tez çalışmasında kullanılmamıştır. Spektrogramda yatay eksen zamanı, dikey eksen frekansı temsil eder ve renk, desibel cinsinden ses gücünün bir boyutu olduğundan dolayı spektrogram üzerinde yapılan herhangi bir veri artırma yöntemi sesin gerçek dünyadaki fiziksel bir dönüşümüne karşılık gelmek zorundadır. Aksi takdirde bu tür yöntemler model performansı üzerinde bozucu etkiye yol açmaktadır. Zhou ve diğerleri (Zhou vd. 2022) tarafından “spektrogram kısıtı” (spectrogram constraint) olarak ifade edilen bu durumu yaptıkları çalışmanın deneysel sonuçları doğrulamaktadır.

Uygulanan segmentasyon işlemi, toplam örnek sayısını artırmak ve gemilerden yayılan gürültü özelliklerinin periyodikliğini ve geçerliliğini bozmadan sabit uzunlukta girdiler elde etmek amacıyla gerçekleştirilmiştir. Bu çalışmada 4 farklı segmentasyon yöntemi uygulanmıştır.

Segment uzunluğunun kısaltılması, her bir sınıftan elde edilen örnek sayısını artırarak derin öğrenme modellerinin veri azlığına karşı daha dayanıklı eğitilmesine katkı sağlamaktadır. Bu durum, derin öğrenme modellerinin veri açlığını gidermeye yardımcı

olabilmekte ve potansiyel olarak aşırı öğrenmeyi azaltabilmektedir. Daha kısa segmentler, modelin daha çeşitli örüntüler görmesini sağlayarak daha iyi genelleme yeteneği kazanmasına ve aşırı öğrenmenin azalmasına katkıda bulunabilir. Ancak, çok kısa segmentler, zaman içinde daha uzun süreler boyunca kendini tekrar eden veya birbiriyle ilişkili olan örüntüleri kaybetme riskini de taşımaktadır. Örneğin, bir geminin motorunun sürekli uğultusu, pervanesinin düzenli vuruşları veya belirli bir makinenin çalışma ritmi gibi karakteristik sesler, ancak yeterince uzun bir zaman dilimi boyunca gözlemlendiğinde anlam kazanan özelliklerdir. Bu durum, sinyalin bütünsel bağlamının kaybolmasına ve pervanesinin dönme hızı gibi periyodik bazı karmaşık özelliklerin yakalanamamasına neden olabilir. Başka bir deyişle, eğer ses kaydı çok kısa parçalara bölünürse, model pervanesinin tam bir dönüş döngüsünü veya motorun karakteristik çalışma ritmini algılayamayabilir. Bu da geminin kimliğini belirlemek için kritik olan bu tür ayırt edici özelliklerin gözden kaçmasına yol açabilir.

Derin öğrenme modellerinin eğitimi esnasında modelin genelleme yeteneğini artırmak ve aşırı öğrenmeyi (overfitting) önlemek amacıyla büyük miktarda etiketli veriye ihtiyaç duyulur. Ancak, su altı akustiği alanında gürültüsüz ve etiketli veri toplamanın maliyetli ve zor olmasından dolayı sınırlı veri setleriyle çalışmak bir zorunluluk olmaktadır. Bu kısıtı aşmak amacıyla literatürde yaygın olarak kullanılan veri artırma (data augmentation) yöntemleri değerlendirilmiştir. Görüntü sınıflandırma problemlerinde sıklıkla başvuru görülen görüntü döndürme (rotation) ve dikey çevirme (vertical flip) gibi dönüşümler, spektrogram tabanlı sınıflandırma problemleri için uygun değildir. Spektrogramlar yatay ekseninde zamanı, dikey ekseninde ise frekansı temsil eder. Bir spektrogramın yatay/dikey ekseninde döndürülmesi gibi işlemler zaman ve frekans eksenlerinden elde edilecek anlamlı veri elde edilmesine engel olur. Bu durum ise akustik sinyalin fiziksel gerçekliğini bozar ve anlamsız bir veri üretilmesine yol açar. Bu nedenle, akustik sinyalin zaman–frekans yapısını bozan geometrik dönüşümler yerine, sinyalin fiziksel doğasıyla uyumlu iki artırma stratejisi benimsenmiştir: ses verisinin farklı uzunluklarda segmentlere ayrılması ve ses sinyallerine Gauss gürültüsü eklenmesi. Gürültü tabanlı bu artırma, spektrogramın zaman–frekans yapısını koruduğundan geometrik dönüşümlerin aksine sinyalin fiziksel anlamını bozmamakta; ilgili gürültü rejimleri ve etkileri Bölüm 4.6'da ele alınmaktadır.

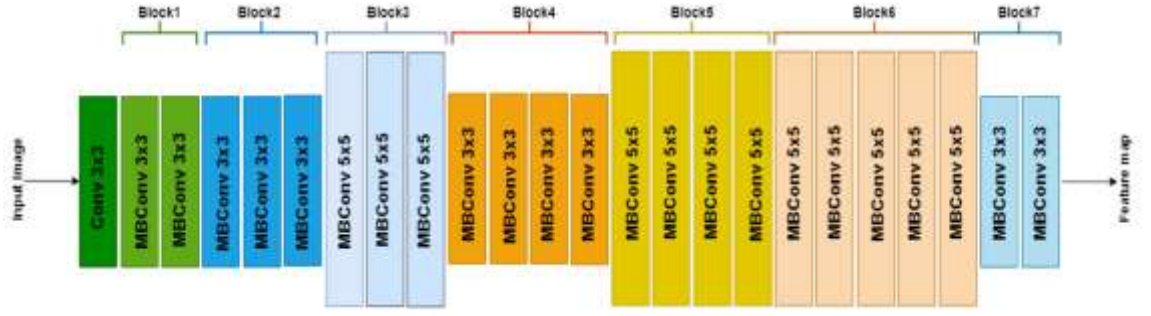
Bu çalışmada 10 saniye, 5 saniye, 3 saniye ve 1 saniyelik olmak üzere dört segmentasyon senaryosu oluşturulmuş; bu senaryolardan sırasıyla 1797, 3644, 6126 ve 18524 spektrogram elde edilmiştir. Segment uzunluğu kısaldıkça elde edilen spektrogram sayısı artmakta; en kısa pencere (1 saniye) en uzun pencereye (10 saniye) kıyasla yaklaşık on kat daha fazla örnek üretmektedir. Segment uzunluğunun bu şekilde denetimli olarak değiştirilmesi, bu değişkenin sınıflandırma başarımı üzerindeki etkisinin Bölüm 4.6’da kurulan faktöriyel deney tasarımı çerçevesinde sistematik biçimde incelenmesine olanak tanımıştır.

#### **4.3 Spektrogram Oluşturma ve Sınıf Dağılımları**

Segmentlere ayrılmış ses dosyaları CNN modeline girdi oluşturmak üzere spektrogramlara dönüştürülmüştür. Spektrogram üretiminde aynı ön işleme adımları korunmuş ve veriler 10 saniye, 5 saniye, 3 saniye ve 1 saniyelik olmak üzere 4 farklı segment uzunluğunda bölütlenerek her segment uzunluğu için ayrı bir spektrogram kümesi elde edilmiştir. Oluşturulan spektrogramlar, gemi büyüklüğüne göre 5 farklı sınıfa ayrılmıştır.

#### **4.4 CNN Model Mimarisi ve Eğitim Detayları**

Bu çalışmada, su altı akustik hedef tanıma performansını artırmak amacıyla, ImageNet veri setiyle eğitilmiş bir evrimsel sinir ağı olan EfficientNetB1 mimarisi transfer öğrenme yöntemiyle kullanılmıştır (Tan ve Le 2019). Şekil 4.7’de EfficientNetB1 modelinin mimarisi gösterilirken Şekil 4.8’de ise modelin parametreleri yer almaktadır.



Şekil 4.7 EfficientNetB1 mimarisi

| Layer (type)                                      | Output Shape        | Param #   |
|---------------------------------------------------|---------------------|-----------|
| input_layer_1 (InputLayer)                        | (None, 224, 224, 3) | 0         |
| sequential (Sequential)                           | (None, 224, 224, 3) | 0         |
| efficientnetb1 (Functional)                       | (None, 7, 7, 1280)  | 6,575,239 |
| global_average_pooling2d (GlobalAveragePooling2D) | (None, 1280)        | 0         |
| dropout (Dropout)                                 | (None, 1280)        | 0         |
| normalization_1 (Normalization)                   | (None, 1280)        | 2,561     |
| dense (Dense)                                     | (None, 5)           | 6,405     |

Total params: 6,584,205 (25.12 MB)  
 Trainable params: 6,519,589 (24.87 MB)  
 Non-trainable params: 64,616 (252.41 KB)

Şekil 4.8 EfficientNetB1 modelinin parametreleri

Modelin eğitim konfigürasyonu, su altı akustik hedef sınıflandırma literatüründeki yaklaşımlar ve hesaplama verimliliği göz önünde bulundurularak yapılandırılmıştır. Modelin eğitiminde Tensorflow, Keras, Scikit-learn ve Numpy kütüphaneleri kullanılmıştır. Modele 224x224x3 boyutlarındaki görüntüler veri girişi olarak verilmiştir (He vd. 2015). Modelde aktivasyon fonksiyonu olarak “softmax” kullanılmış olup “pooling” değeri “average” olarak ele alınmıştır (Hong vd. 2021, Tian vd. 2021, Yao vd. 2023). Modelde “batch size” değeri 32, “epoch” değeri ise 100 olarak kullanılmıştır. Modelde “optimizer” olarak, “Adam” kullanılmıştır (Hong vd. 2021, Tian vd. 2021, Zeng

vd. 2025). “loss” fonksiyonu olarak ise “categorical crossentropy” kullanılmıştır (Tian vd. 2021, Xu vd. 2024, Zeng vd. 2025).

Ağ yapısına doğrusal olmayan bir karakter kazandırmak ve veri kümesindeki karmaşık örüntülerin öğrenilmesini sağlamak amacıyla ReLU (Rectified Linear Unit) aktivasyon fonksiyonu tercih edilmiştir (Hong vd. 2021, Neupane ve Seok 2020, Xu vd. 2024, Zeng vd. 2025). Modelin ezberleme (overfitting) eğilimini minimize etmek ve genelleme kabiliyetini artırmak adına düzenlileştirme (regularization) tekniği olarak dropout yöntemi mimariye dahil edilmiştir (Du ve Hong 2024, Hong vd. 2021, Neupane ve Seok 2020). Bu doğrultuda, çıkış katmanından önce nöronların rastgele %10'u geçici olarak devre dışı bırakılarak ağın belirli örüntülere bağımlı kalması engellenmiştir (Du ve Hong 2024).

Çok sınıflı sınıflandırma kapsamında, ağ mimarisinin çıkış katmanı çalışmadaki 5 sınıfla uyumlu olacak şekilde 5 nöronlu olarak yapılandırılmıştır (Hong vd. 2021, Santos-Domínguez vd. 2016, Yang vd. 2024). Bu katmanda, karar mekanizmasının ham çıktı değerlerini olasılıksal bir ifadeye dönüştürmek amacıyla “softmax” aktivasyon fonksiyonu kullanılmıştır (He vd. 2015, Hong vd. 2021, Lee vd. 2020, Shen vd. 2018, Xu vd. 2024).

Modelin eğitim sürecinin kararlılığını sağlamak ve ezberleme riskini en aza indirmek amacıyla, sürece dinamik öğrenme oranı ve geri çağırma fonksiyonları (callbacks) dahil edilmiştir (Doan vd. 2022, He vd. 2015). Başlangıç öğrenme oranı, eğitim esnasında sabit bırakılmamış, doğrulama kaybının (validation loss) gidişatına göre adaptif olarak ayarlanmıştır. Bu doğrultuda kullanılan dinamik öğrenme oranı azaltma fonksiyonu, doğrulama kaybının ardışık 5 epoch boyunca iyileşme göstermediği durağanlık (plateau) evrelerinde, öğrenme oranını 0,1 azaltma katsayısı ile çarparak 10 kat küçültecek şekilde yapılandırılmıştır. Ulaşılabilecek en düşük öğrenme oranı  $10^{-6}$  olarak belirlenmiştir. Modelin gereksiz hesaplama yapmasını önlemek ve ezberlemeye geçişini engellemek amacıyla erken durdurma (early stopping) fonksiyonu kullanılmıştır. Doğrulama kaybında ardışık 10 epoch boyunca (patience) herhangi bir gelişim olmadığı takdirde eğitim süreci otomatik olarak sonlandırılmaktadır (Hong vd. 2021, Zeng vd. 2025). Erken

durdurma mekanizması tetiklendiğinde, modelin genellenebilirlik kapasitesini en üst seviyede tutmak adına, modelin parametreleri doğrulama kaybının en düşük olduğu epochtaki değerlerine geri döndürülmektedir.

Veri kümesinin sınırlı boyutu ve sınıflar arasındaki dengesiz dağılım dikkate alınarak tabakalı 10 katlı çapraz doğrulama (stratified 10-fold cross-validation) yöntemi uygulanmıştır. Bu yöntem sınıf oranlarının her bir katlamada (fold) sabit tutulmasıyla, sınıf dengesizliğinden kaynaklanabilecek olası yanlılıkların (bias) en aza indirgenmesi amacıyla tercih edilmiştir (Zhou vd. 2022).

Çapraz doğrulamanın her katında veri setinin %10'u test kümesi olarak ayrılmış; kalan %90'lık kısım ise %80/%20 oranında eğitim (toplamın %72'si) ve doğrulama (%18'i) kümelerine bölünmüştür.

Model, her bir eğitim senaryosunda (segment uzunluğu ve gürültü rejimi kombinasyonlarında) maksimum 100 epoch boyunca eğitilmiştir.

#### **4.5 Performans Değerlendirme Kriterleri**

Geliştirilen su altı akustik hedef sınıflandırma modelinin etkinliğini, güvenilirliğini ve pratik uygulanabilirliğini tespit etmek, yalnızca genel bir doğruluk oranı belirtmenin ötesinde, çok yönlü ve titiz bir değerlendirme sürecini zorunlu kılmaktadır. Özellikle savunma sanayii gibi hata maliyetinin yüksek olduğu kritik alanlarda, bir modelin sadece ne kadar başarılı olduğu değil, aynı zamanda ne tür hatalar yaptığı ve bu hataların hangi sınıflar arasında yoğunlaştığı kritik önem taşır. Bu bağlamda, modelin performansını bütüncül bir bakış açısıyla analiz etmek ve zayıf yönlerini ortaya çıkarmak amacıyla, makine öğrenmesi modellerinin performansını anlamakta kullanılan hata matrisinden (confusion matrix) ve bu matristen türetilen metriklerden faydalanılmıştır. Bu kriterler, modelin sınıflandırma kabiliyetinin derinlemesine anlaşılması için temel bir çerçeve sunmaktadır.

#### 4.5.1 Hata matrisi

Hata matrisi, bir sınıflandırma modelinin performansını görselleştiren ve her bir sınıf için doğru ve yanlış tahmin sayılarını detaylı bir şekilde özetleyen güçlü bir analiz aracıdır. Hata matrisi, modelin genel doğruluk oranının ardındaki karmaşık dinamikleri anlamak için vazgeçilmezdir. Özellikle, modelin hangi hedef sınıflarını (örneğin, bir askeri gemi ile bir sivil konteyner gemisi) birbirleriyle karıştırdığını net bir şekilde göstererek, potansiyel riskleri ve modelin iyileştirilmesi gereken alanları doğrudan işaret eder.

Hata matrisinin yorumlanmasında kullanılan temel bileşenler şunlardır:

- **Doğru Pozitif (DP):** Modelin bir örneği doğru bir şekilde ilgili sınıfa ait olarak tahmin ettiği durumdur. Örneğin, gerçekte "Class A" olan bir geminin, model tarafından da "Class A" olarak tahmin edilmesi.
- **Doğru Negatif (DN):** Modelin bir örneği doğru bir şekilde ilgili sınıfa ait olmadığını tahmin ettiği durumdur. Örneğin, gerçekte "Class B" olan bir geminin, model tarafından "Class A" olarak tahmin edilmemesi (yani, "Class A" dışındaki bir sınıf olarak tahmin edilmesi).
- **Yanlış Pozitif (YP):** Modelin bir örneği yanlış bir şekilde ilgili sınıfa ait olarak tahmin ettiği durumdur (Tip I hata). Örneğin, gerçekte "Class B" olan bir geminin, model tarafından yanlışlıkla "Class A" olarak tahmin edilmesi.
- **Yanlış Negatif (YN):** Modelin bir örneği yanlış bir şekilde ilgili sınıfa ait olmadığını tahmin ettiği durumdur (Tip II hata). Örneğin, gerçekte "Class A" olan bir geminin, model tarafından yanlışlıkla "Class A" dışında bir sınıf olarak tahmin edilmesi.

#### 4.5.2 Performans metrikleri

Hata matrisinden elde edilen DP, DN, YP ve YN değerleri, modelin performansını standart, objektif ve karşılaştırılabilir ölçütlerle ifade etmek için kullanılır. Bu metrikler, modelin farklı operasyonel senaryolardaki davranışını anlamamızı sağlar.

Doğruluk (Accuracy), modelin tüm tahminleri içinde doğru yaptığı tahminlerin oranıdır. Genel bir performans göstergesi olmasına rağmen, veri setindeki sınıf dağılımı dengesiz olduğunda (bazı sınıfların diğerlerinden çok daha fazla örneğe sahip olması) yanıltıcı olabilir. Bu nedenle, tek başına yeterli bir ölçüt değildir.

$$\text{Doğruluk} = \frac{DP + DN}{DP + DN + YP + YN}$$

Kesinlik (Precision), modelin pozitif olarak etiketlediği hedefler arasında ne kadarının gerçekten pozitif olduğunu ölçer. Bu metrik, yanlış pozitiflerin maliyetinin yüksek olduğu durumlar için kritiktir. Örneğin, zararsız bir sivil geminin tehdit olarak etiketlenmesi (düşük kesinlik) gereksiz yere alarm durumuna geçilmesine neden olabilir.

$$\text{Kesinlik} = \frac{DP}{DP + YP}$$

Duyarlılık (Recall), gerçekte pozitif olan hedeflerin ne kadarının model tarafından başarıyla tespit edildiğini gösterir. Bu metrik, yanlış negatiflerin maliyetinin çok daha yüksek olduğu savunma uygulamaları için hayati öneme sahiptir. Gerçek bir tehdidin tespit edilememesi (düşük duyarlılık), feci sonuçlar doğurabilir.

$$\text{Duyarlılık} = \frac{DP}{DP + YN}$$

F1-Skoru (F1-Score), kesinlik ve duyarlılık metriklerinin harmonik ortalamasıdır. Bu iki metrik arasında bir denge kurar ve özellikle sınıf dağılımının dengesiz olduğu veya hem yanlış pozitif hem de yanlış negatif maliyetlerinin dikkate alınması gereken durumlarda, model performansının daha adil ve güvenilir bir değerlendirmesini sunar.

$$F1 \text{ Skoru} = \frac{2 \cdot \text{Kesinlik} \cdot \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}}$$

Destek (Support) değeri, belirli bir sınıfa ait örneklerin toplam sayısıdır. Bu değer her bir sınıfın veri setindeki dağılımını gösterir ve sınıf dengesizliği hakkında bilgi verir.

Makro ortalama (Macro average), her bir sınıf için metrik (kesinlik, duyarlılık, F1-Skoru) ayrı ayrı hesaplanır ve ardından bu değerlerin basit (ağırlıksız) ortalaması alınarak bulunur. Makro ortalama, tüm sınıflara eşit önem verir. Bu nedenle, azınlık sınıfların

performansındaki düşüşlere karşı hassastır ve dengesiz veri setlerinde modelin genel performansını daha adil bir şekilde yansıtabilir.

Ağırlıklı ortalama (Weighted average), her bir sınıf için metrik (kesinlik, duyarlılık, F1-Skoru) ayrı ayrı hesaplanır ve ardından bu değerler, ilgili sınıfın destek (örnek) sayısına göre ağırlıklandırılarak ortalaması alınarak bulunur. Ağırlıklı ortalama, sınıf dengesizliğini dikkate alır ve daha fazla örneğe sahip sınıflara daha fazla önem verir. Bu, modelin genel performansının veri setindeki gerçek sınıf dağılımını daha iyi yansıtmasını sağlar.

Bu kapsamlı değerlendirme kriterleri seti, geliştirilen CNN modelinin sadece genel başarısını değil, aynı zamanda operasyonel açıdan kritik olan nüanslarını da ortaya koymaktadır.

#### 4.6 Deneysel Sonuçlar ve Performans Analizi

Bu bölümde, ShipsEar veri seti üzerinde EfficientNetB1 mimarisi ile gerçekleştirilen sınıflandırma deneylerinin sonuçları sunulmaktadır. Çalışma, iki bağımsız tasarım değişkeninin sistematik biçimde birlikte ele alındığı faktöriyel bir deneysel tasarım temelinde yapılandırılmıştır. Birinci değişken, ses sinyallerinin bölütlendiği segment uzunluğu olup 10, 5, 3 ve 1 saniye olmak üzere dört düzeyde incelenmiştir. İkinci değişken, eğitim setine uygulanan Gauss gürültüsü artırımının şiddetini belirleyen gürültü rejimidir; söz konusu rejim, segment başına çekilen gürültü seviyesinin dağılım ortalaması bakımından iki düzeyde ele alınmıştır: 0 dB ortalama ve 5 dB standart sapmaya sahip  $N(0, 5^2)$  rejimi ile 3 dB ortalama ve 5 dB standart sapmaya sahip  $N(3, 5^2)$  rejimi. Bu tasarım sonucunda sekiz ana deney elde edilmiştir. Söz konusu sekiz deneye ek olarak, gürültü artırımının başarıma katkısını yalıtılabilmek amacıyla 5 saniyelik segmentasyon için gürültü artırımının uygulanmadığı bir kontrol deneyi de yürütülmüştür.

İki gürültü rejimi yalnızca eklenen gürültü seviyesinin dağılım ortalaması bakımından birbirinden ayrılmaktadır:  $N(3, 5^2)$  rejimi,  $N(0, 5^2)$  rejimine kıyasla ortalama 3 dB daha

yüksek bir gürültü seviyesi eklemekte ve dolayısıyla eğitim örneklerinde ortalama olarak daha düşük bir etkin sinyal-gürültü oranına (SNR) karşılık gelmektedir. Bu yönüyle  $N(3, 5^2)$  rejimi, modelin gürültü dayanıklılığını daha zorlayıcı koşullar altında sınavan bir nitelik taşımaktadır. Tüm deneylerde eğitim hiperparametreleri, ön işleme adımları ve tabakalı 10 katlı çapraz doğrulama protokolü sabit tutulmuş; böylece gözlenen başarımlar farklılıklarının yalnızca segment uzunluğu ve gürültü rejimi ile ilişkilendirilebilmesi hedeflenmiştir. Tüm deneyler, gemi büyüklüğüne göre oluşturulan Class A, Class B, Class C, Class D ve Class E olmak üzere beş sınıf üzerinde gerçekleştirilmiştir.

Sekiz ana deneyin özet başarımlar göstergeleri (ortalama doğrulama doğruluğu ve makro ortalama F1-Skoru, 10 kat üzerinden hesaplanan standart sapmalarıyla birlikte) Çizelge 4.2'de topluca verilmiştir. Alt bölümlerde her segment uzunluğu, iki gürültü rejimi altında ayrı ayrı ele alınmış; öğrenme eğrileri, hata matrisleri ve sınıf bazlı metrikler ayrıntılı biçimde değerlendirilmiştir. Bölüm sonunda ise segment uzunluğu ve gürültü rejiminin etkileri karşılaştırmalı olarak sentezlenmektedir.

Çizelge 4.2 Sekiz ana deneyin ve gürültüsüz kontrol deneyinin özet başarımlar göstergeleri

| <b>Segment Uzunluğu Gürültü Rejimi Doğrulama Doğruluğu F1-Skoru (Makro)</b> |             |                      |                     |
|-----------------------------------------------------------------------------|-------------|----------------------|---------------------|
| 10 saniye                                                                   | $N(0, 5^2)$ | $\%85,92 \pm \%3,14$ | $0,8568 \pm 0,0308$ |
| 10 saniye                                                                   | $N(3, 5^2)$ | $\%77,96 \pm \%3,08$ | $0,7804 \pm 0,0323$ |
| 5 saniye                                                                    | $N(0, 5^2)$ | $\%86,69 \pm \%1,47$ | $0,8655 \pm 0,0143$ |
| 5 saniye                                                                    | $N(3, 5^2)$ | $\%79,67 \pm \%2,78$ | $0,7967 \pm 0,0271$ |
| 5 saniye                                                                    | Gürültü yok | $\%82,37 \pm \%3,21$ | $0,8243 \pm 0,0297$ |
| 3 saniye                                                                    | $N(0, 5^2)$ | $\%88,52 \pm \%1,39$ | $0,8887 \pm 0,0131$ |
| 3 saniye                                                                    | $N(3, 5^2)$ | $\%80,51 \pm \%1,77$ | $0,8084 \pm 0,0199$ |
| 1 saniye                                                                    | $N(0, 5^2)$ | $\%86,02 \pm \%0,72$ | $0,8636 \pm 0,0083$ |
| 1 saniye                                                                    | $N(3, 5^2)$ | $\%78,62 \pm \%1,52$ | $0,7903 \pm 0,0149$ |

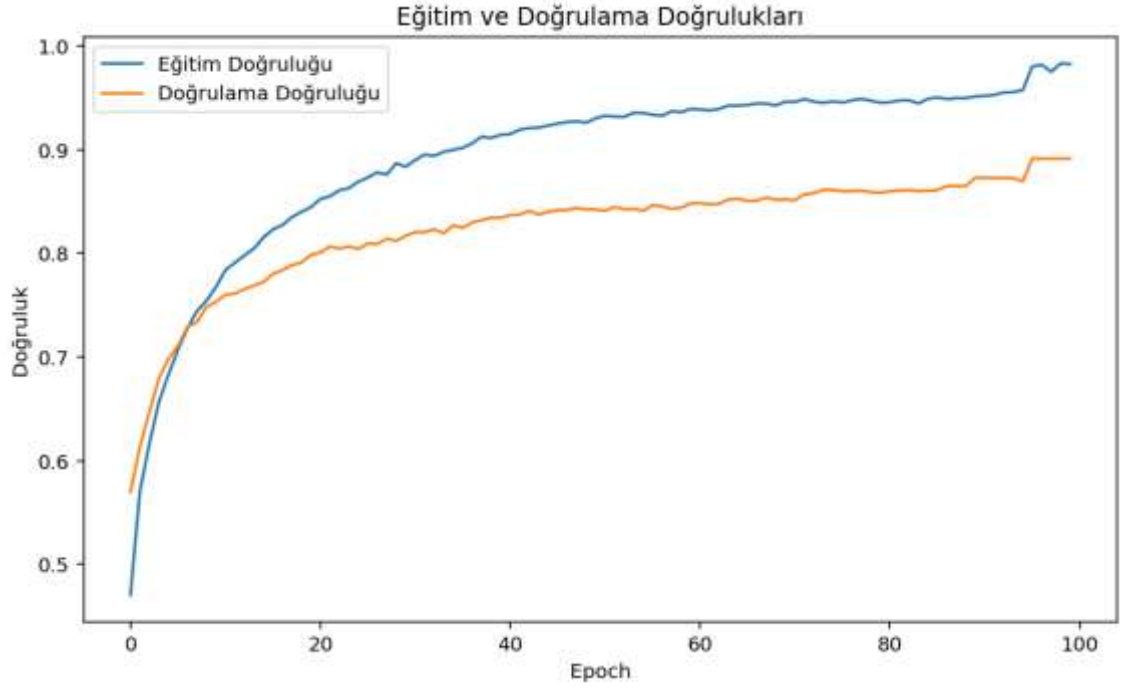
Çizelge 4.2 iki temel örüntüyü ortaya koymaktadır. Birincisi, her iki gürültü rejiminde de doğrulama doğruluğu segment uzunluğuyla tekdüze biçimde değişmemekte; 3 saniyelik

konfigürasyonda tepe noktasına ulaşmaktadır. İkincisi,  $N(3, 5^2)$  rejimi tüm segment uzunluklarında  $N(0, 5^2)$  rejiminin belirgin biçimde altında kalmakta ve iki rejim arasındaki fark yaklaşık 7-8 puan bandında seyretmektedir. Bu iki örüntü, sırasıyla segment uzunluğunun ve gürültü rejiminin model başarımı üzerindeki bağımsız etkilerini işaret etmektedir.

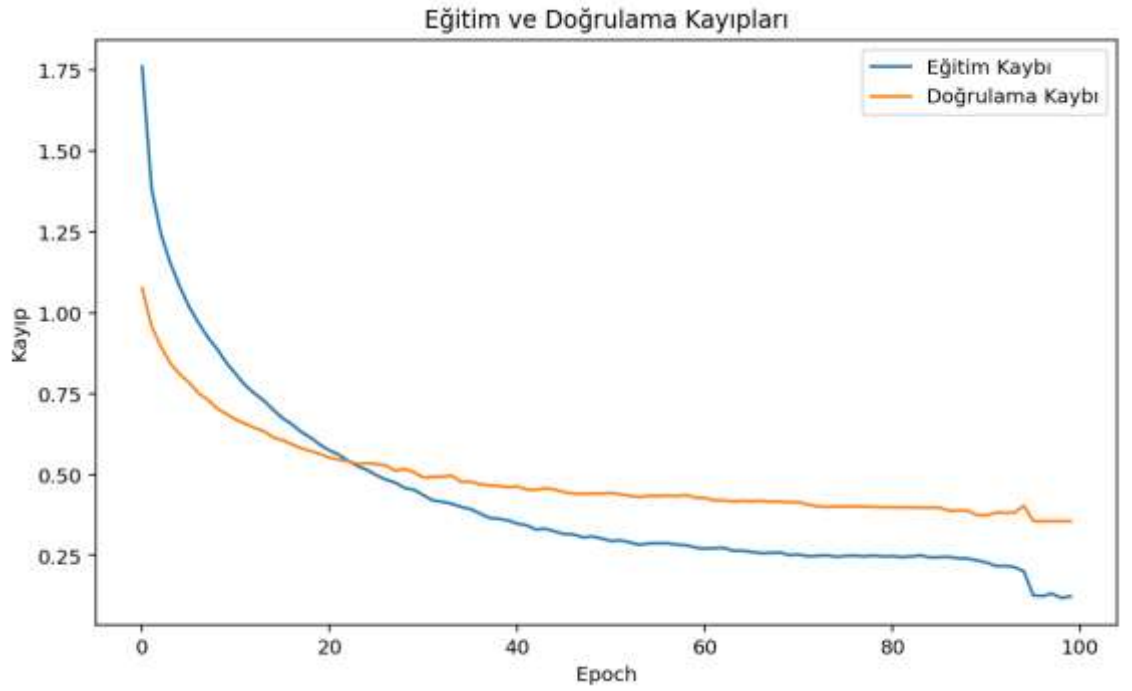
#### 4.6.1 10 saniyelik segmentasyon deneyi sonuçları

10 saniyelik segmentasyon, ses kayıtlarından toplam 1797 spektrogram üretmekte ve dört segment uzunluğu arasında en düşük örneklem hacmini oluşturmaktadır. Bu alt bölümde söz konusu konfigürasyon,  $N(0, 5^2)$  ve  $N(3, 5^2)$  gürültü rejimleri altında ayrı ayrı ele alınmaktadır.

$N(0, 5^2)$  rejiminde ortalama doğrulama doğruluğu  $\%85,92 \pm \%3,14$  ve makro F1-Skoru  $0,8568 \pm 0,0308$  olarak elde edilmiştir. Şekil 4.9 ve Şekil 4.10'da verilen öğrenme eğrileri, eğitim doğruluğunun erken epoch'larda 0,98 düzeyine yaklaşırken doğrulama doğruluğunun 0,86–0,89 bandında dengelendiğini göstermektedir. Eğitim ve doğrulama eğrileri arasındaki bu açıklık, sınırlı örneklem hacminin (yalnızca 90 ham kayıttan üretilen 1797 örnek) modelde belirli düzeyde bir ezberleme (overfitting) eğilimine yol açtığına işaret etmektedir. Nitekim standart sapmanın dört  $N(0, 5^2)$  deneyi arasındaki en yüksek değere ( $\pm \%3,14$ ) ulaşması da bu sınırlı örneklem hacminin katlamalar arası kararlılığı düşürdüğünü desteklemektedir.

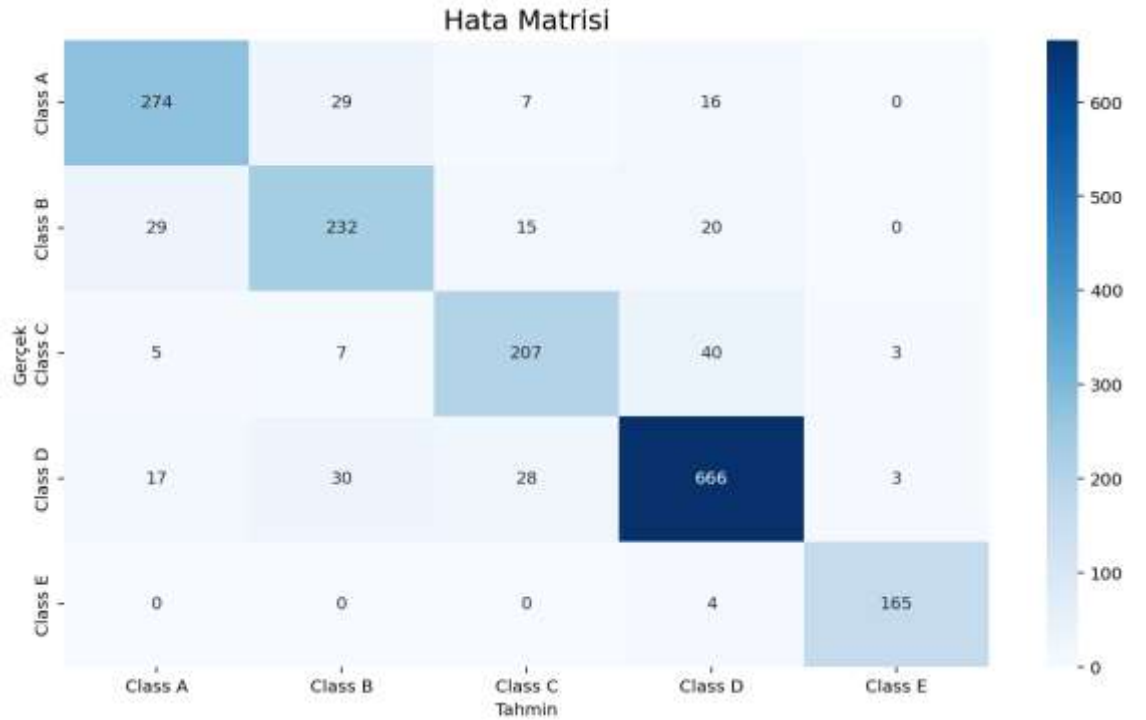


Şekil 4.9 10 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu eğitim ve doğrulama doğrulukları

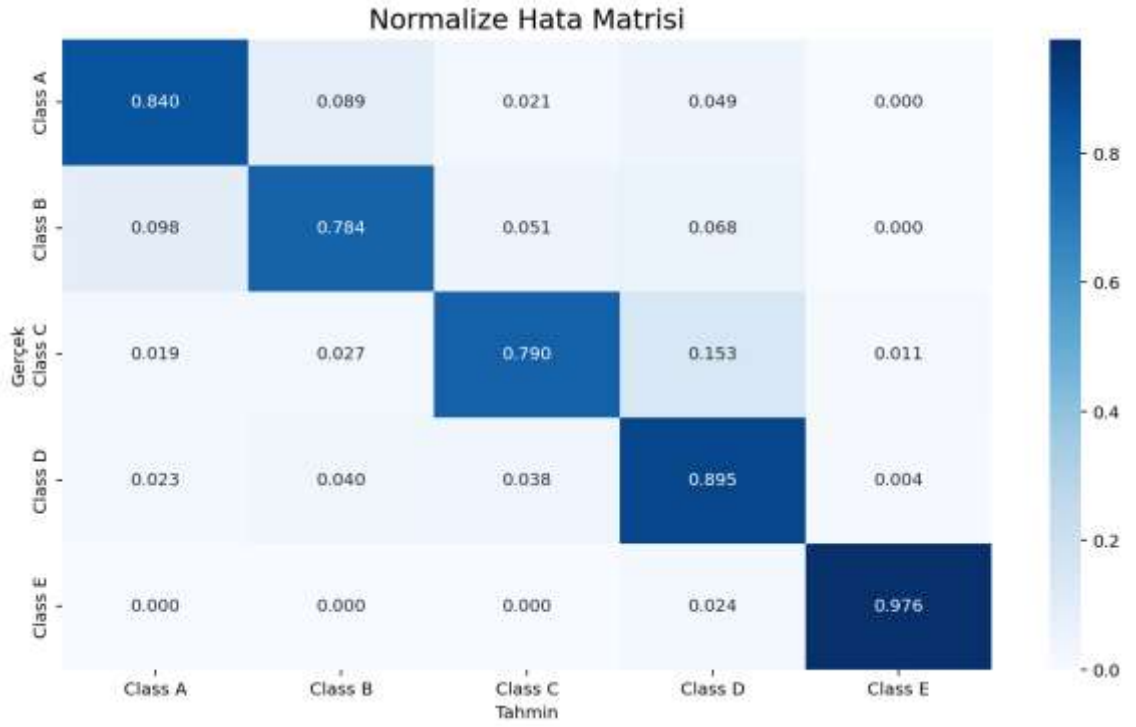


Şekil 4.10 10 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu eğitim ve doğrulama kayıpları

Şekil 4.11 ve Şekil 4.12'deki hata matrisleri ile Çizelge 4.3'teki sınıf bazlı metrikler birlikte değerlendirildiğinde, modelin Class E (F1-Skoru = 0,9706) ve Class D (F1-Skoru = 0,8940) sınıflarını yüksek bir ayırt edilebilirlikle sınıflandırdığı görülmektedir. Class E'nin en yüksek başarımı, doğal ortam gürültüsünün gemi kaynaklı sinyallere kıyasla belirgin biçimde farklı bir yapıya sahip olmasıyla; tonal hat veya sistematik harmonik bileşen içermeyen, geniş bantlı ve stokastik bir zaman-frekans dağılımı sergilemesiyle açıklanabilir (Zhu vd. 2022). Buna karşılık Class A, B ve C, spektral içeriklerinin özellikle 0–4 kHz düşük-orta frekans bandında örtüşmesi nedeniyle kendi aralarında belirgin bir karışma (confusion) örüntüsü göstermekte; bu sınıfların F1-Skorları 0,7811 ile 0,8418 aralığında yer almaktadır.



Şekil 4.11 10 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu hata matrisi



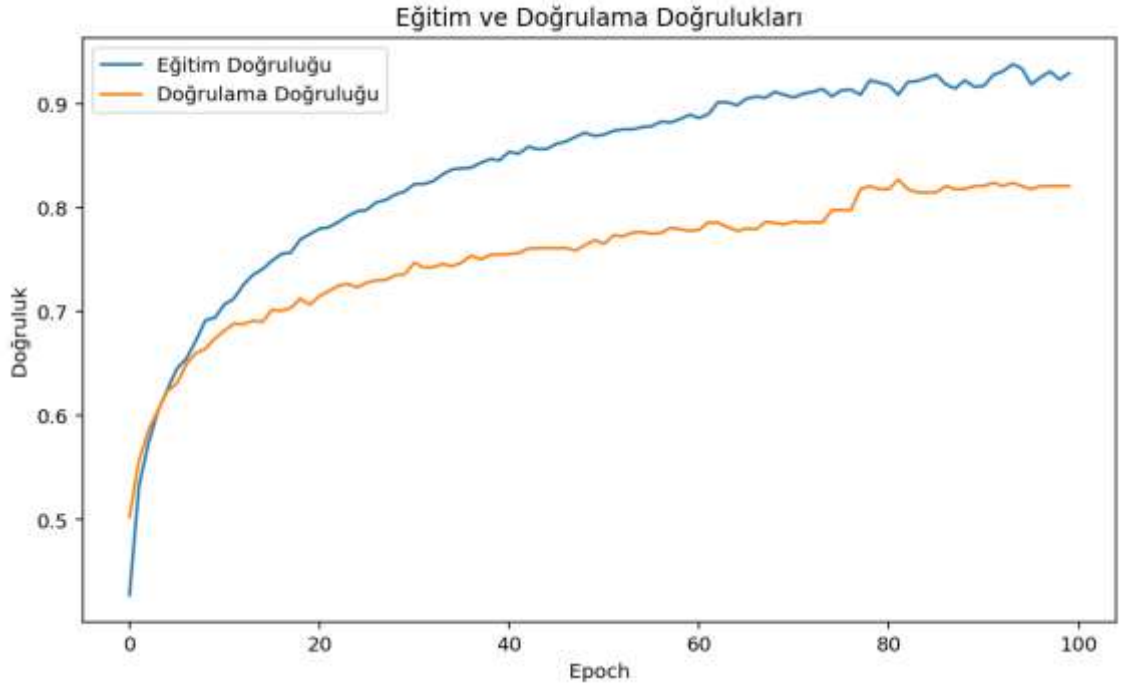
Şekil 4.12 10 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu normalize hata matrisi

Çizelge 4.3 10 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu performans metrikleri

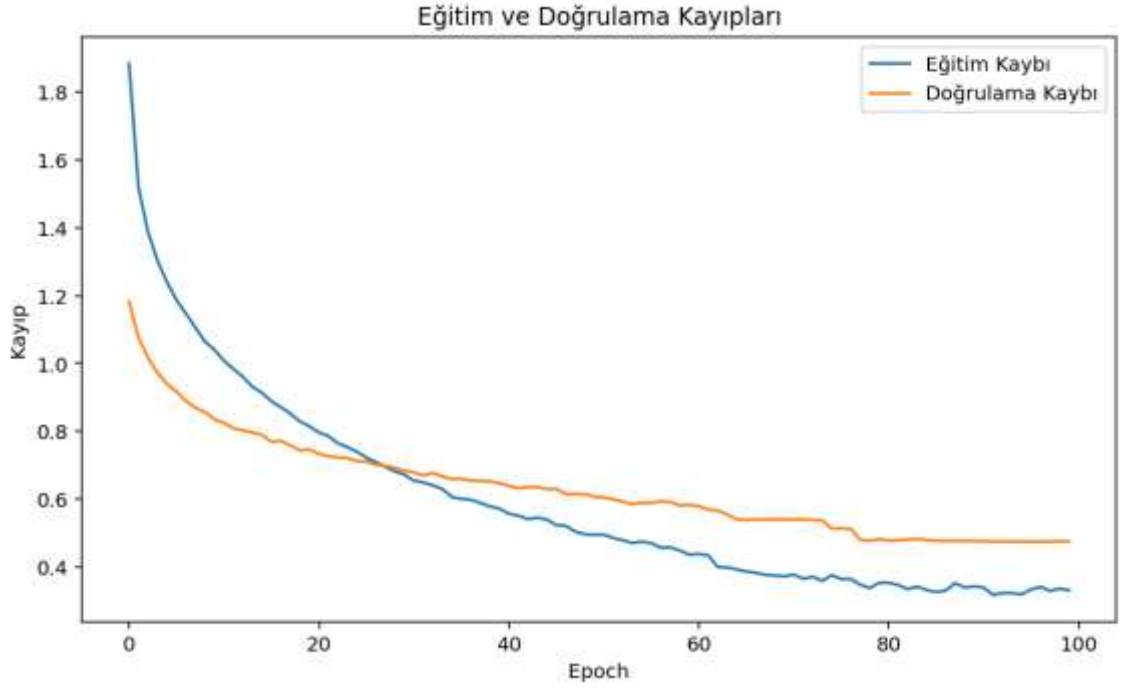
| Sınıflar              | Kesinlik | Duyarlılık | F1-Skoru | Destek |
|-----------------------|----------|------------|----------|--------|
| Class A               | 0,8431   | 0,8405     | 0,8418   | 326    |
| Class B               | 0,7785   | 0,7838     | 0,7811   | 296    |
| Class C               | 0,8054   | 0,7901     | 0,7977   | 262    |
| Class D               | 0,8928   | 0,8952     | 0,8940   | 744    |
| Class E               | 0,9649   | 0,9763     | 0,9706   | 169    |
| <b>Doğruluk</b>       |          |            | 0,8592   | 1797   |
| <b>Makro Ort.</b>     | 0,8569   | 0,8572     | 0,8570   | 1797   |
| <b>Ağırlıklı Ort.</b> | 0,8590   | 0,8592     | 0,8591   | 1797   |

$N(3, 5^2)$  rejiminde gürültü seviyesinin ortalamada 3 dB yükseltilmesiyle birlikte ortalama doğrulama doğruluğu  $77,96 \pm 3,08$  düzeyine gerilemiş; makro F1-Skoru ise  $0,7804 \pm$

0,0323 olarak elde edilmiştir.  $N(0, 5^2)$  rejimine kıyasla ortalama doğrulukta 7,96 puanlık belirgin bir düşüş gözlenmektedir. Şekil 4.13 ve Şekil 4.14'teki öğrenme eğrileri, doğrulama doğruluğunun bu rejimde daha düşük bir platoda (0,74–0,83) dengelendiğini ve eğitim–doğrulama açıklığının korunduğunu göstermektedir. Bu bulgu, artan gürültü şiddetinin sınıf ayırıcı bilgiyi kısmen örttüğüne ve öğrenme problemini zorlaştırdığına işaret etmektedir.

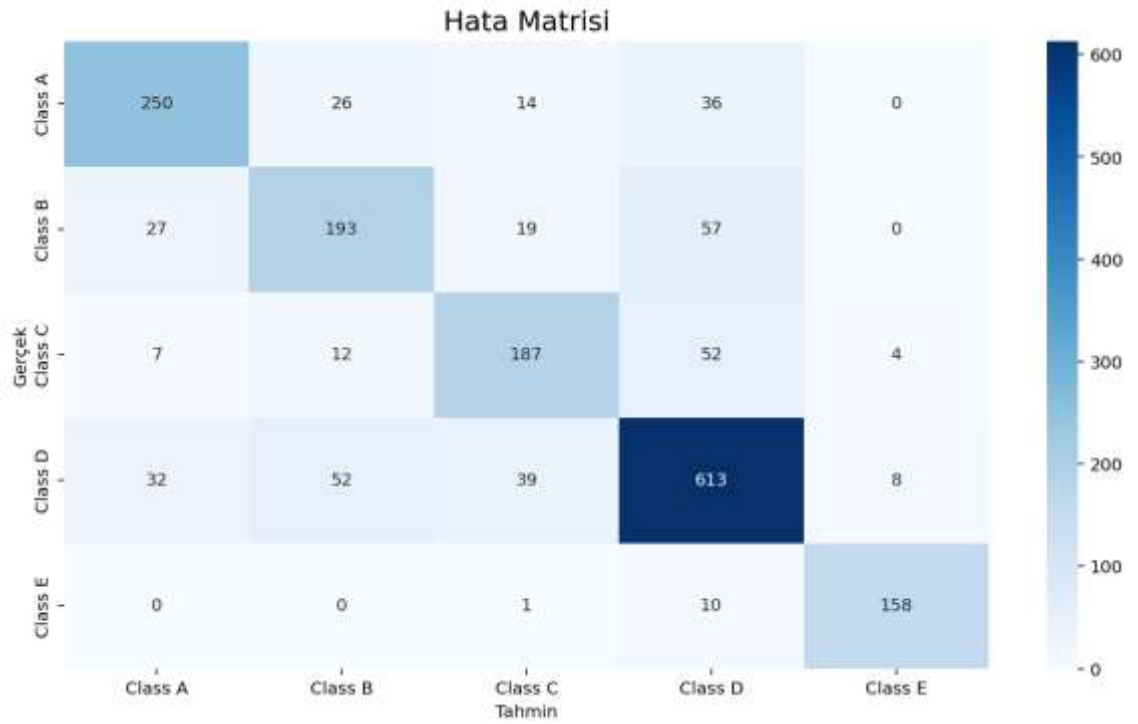


Şekil 4.13 10 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu eğitim ve doğrulama doğrulukları

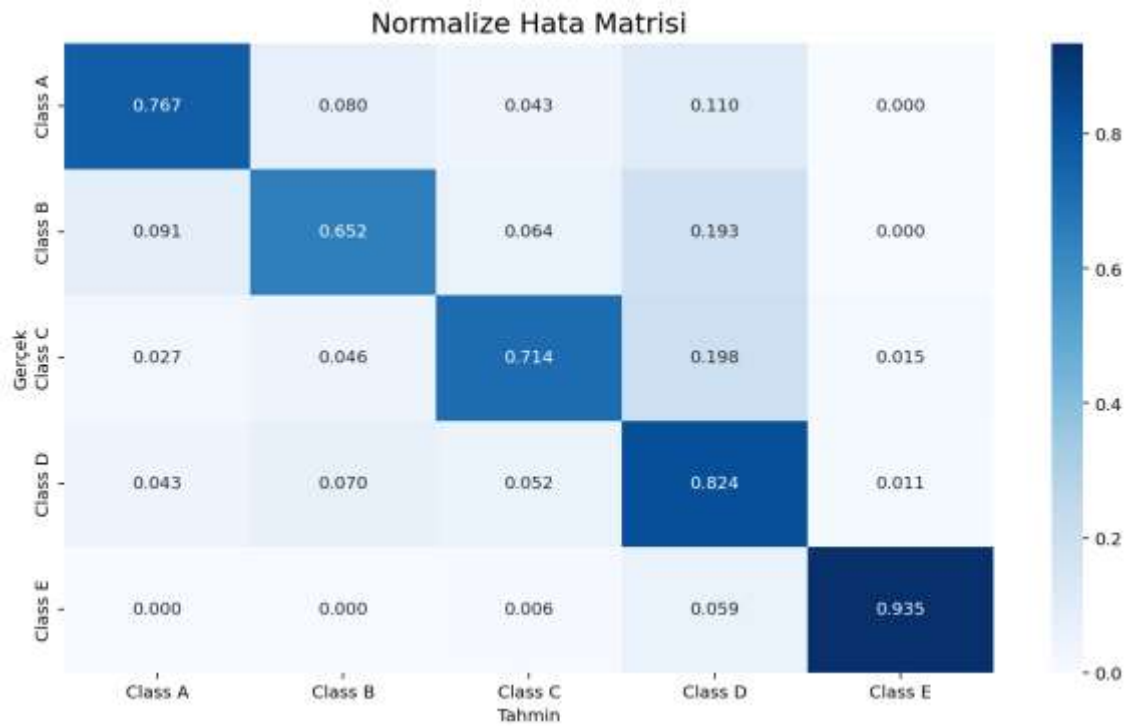


Şekil 4.14 10 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu eğitim ve doğrulama kayıpları

Çizelge 4.4'te görüldüğü üzere, gürültü şiddetindeki artıştan en fazla etkilenen sınıflar, Class B ( $F1 = 0,6667$ ) ve Class C ( $F1 = 0,7165$ ) olmuştur. Bu sınıfların akustik imzaları görece zayıf ve belirsiz olduğundan, eklenen gürültü bu sınıfları ayırt edilebilirlik eşiğinin altına taşımaktadır. Buna karşılık Class E,  $F1$ -Skorunu 0,9322 düzeyinde koruyarak gürültüye en dayanıklı sınıf olma özelliğini sürdürmektedir.



Şekil 4.15 10 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu hata matrisi



Şekil 4.16 10 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu normalize hata matrisi

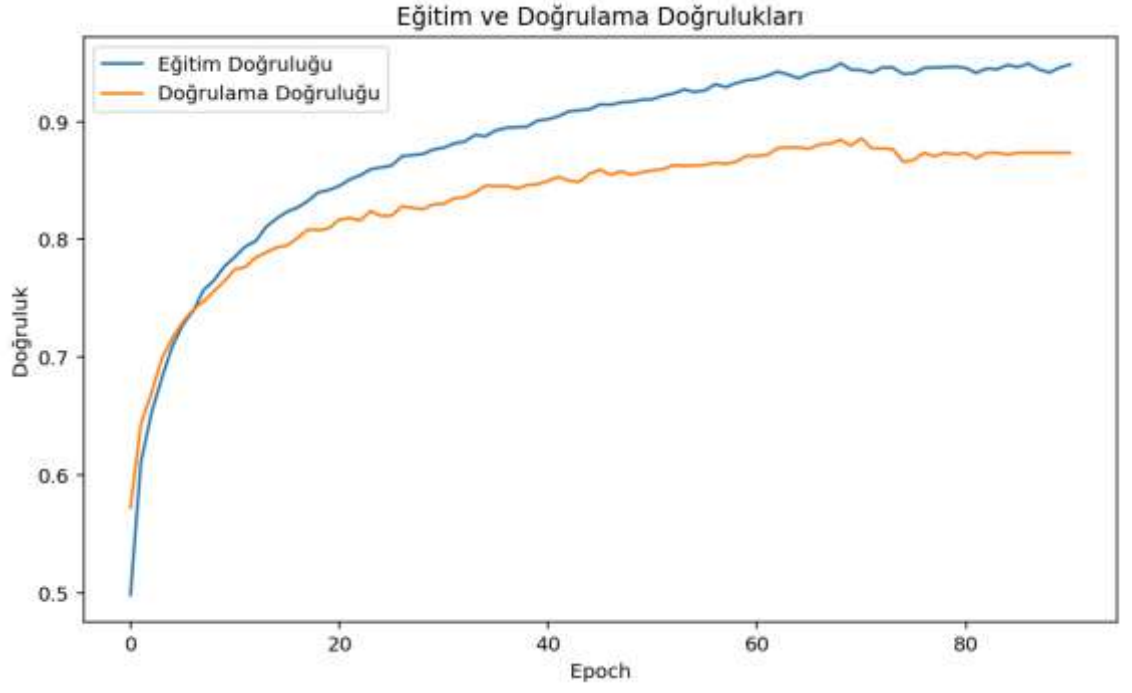
Çizelge 4.4 10 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu performans metrikleri

| Sınıflar              | Kesinlik | Duyarlılık | F1-Skoru | Destek |
|-----------------------|----------|------------|----------|--------|
| Class A               | 0,7911   | 0,7669     | 0,7788   | 326    |
| Class B               | 0,6820   | 0,6520     | 0,6667   | 296    |
| Class C               | 0,7192   | 0,7137     | 0,7165   | 262    |
| Class D               | 0,7982   | 0,8239     | 0,8108   | 744    |
| Class E               | 0,9294   | 0,9349     | 0,9322   | 169    |
| <b>Doğruluk</b>       |          |            | 0,7796   | 1797   |
| <b>Makro Ort.</b>     | 0,7840   | 0,7783     | 0,7810   | 1797   |
| <b>Ağırlıklı Ort.</b> | 0,7786   | 0,7796     | 0,7789   | 1797   |

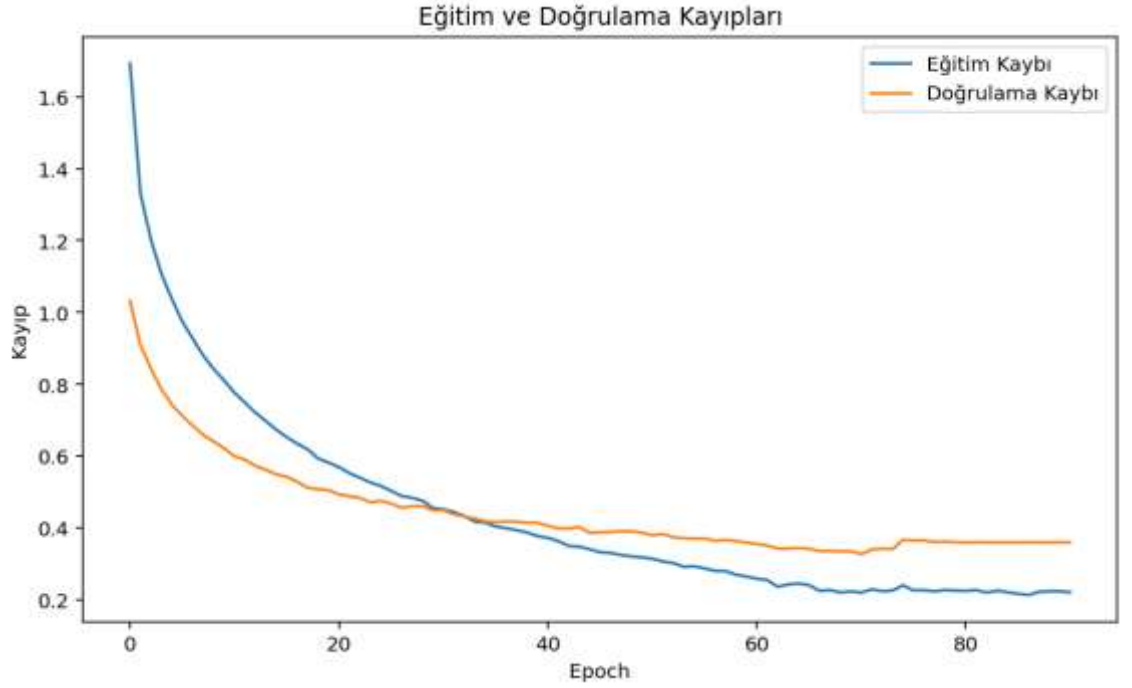
#### 4.6.2 5 saniyelik segmentasyon deneyi sonuçları

5 saniyelik segmentasyon, örnek havuzunu 3644 spektrograma çıkararak veri hacmini yaklaşık iki katına yükseltmiştir. Sınıflar arası oransal dağılımın bu deneyde de korunduğu görülmektedir.

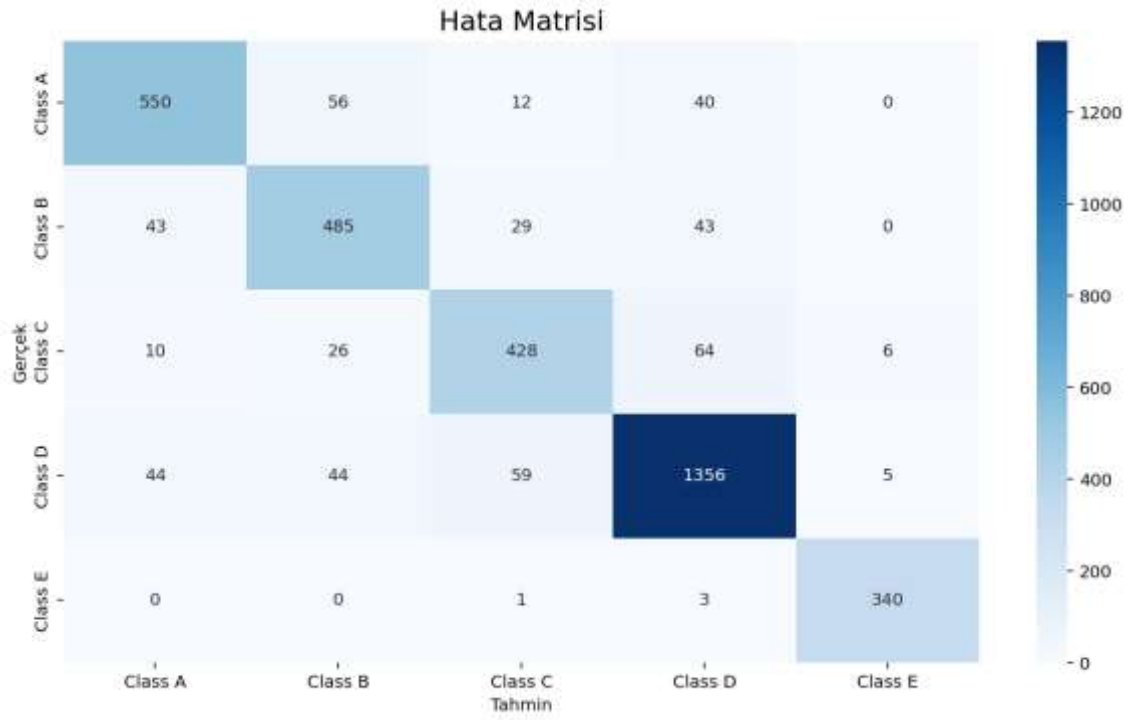
$N(0, 5^2)$  rejiminde ortalama doğrulama doğruluğu  $\%86,69 \pm \%1,47$ , makro F1-Skoru ise  $0,8655 \pm 0,0143$  olarak elde edilmiştir. 10 saniyelik konfigürasyona kıyasla ortalama doğrulukta 0,77 puanlık bir iyileşme ve standart sapmada belirgin bir azalma ( $\pm\%3,14 \rightarrow \pm\%1,47$ ) gözlenmektedir. Şekil 4.17 ve Şekil 4.18'deki öğrenme eğrileri, artan örneklem hacminin katlamalar arası kararlılığı artırdığını ve doğrulama doğruluğunun daha dar bir aralıkta dengelendiğini göstermektedir.



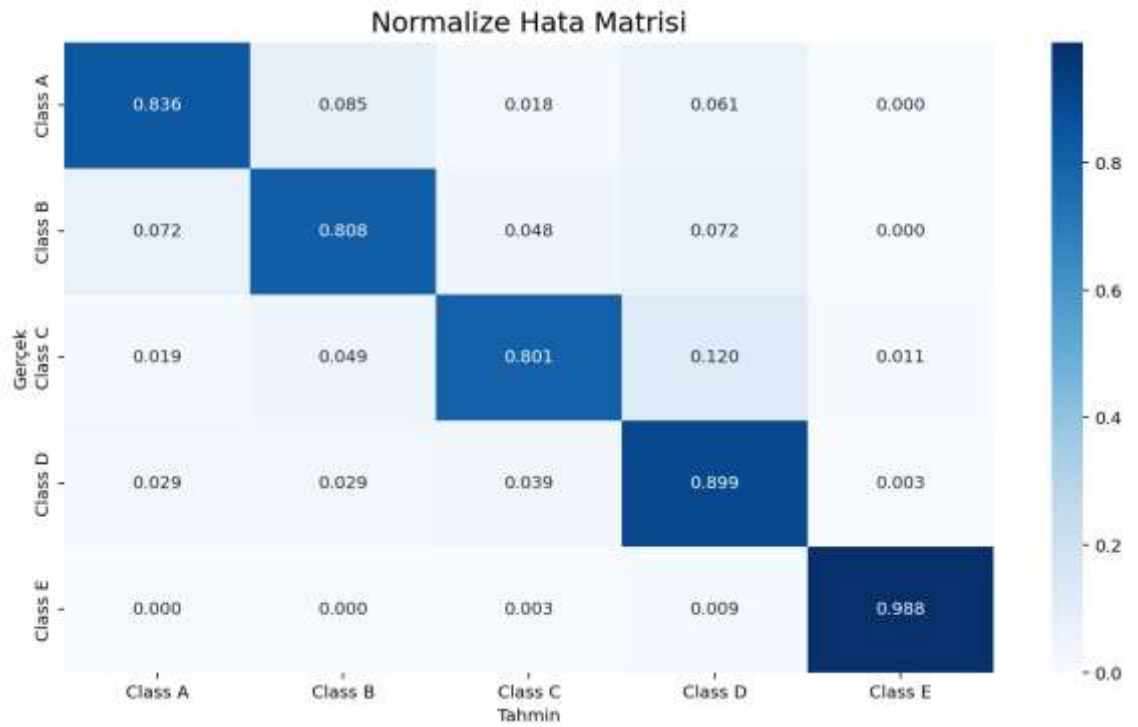
Şekil 4.17 5 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu eğitim ve doğrulama doğrulukları



Şekil 4.18 5 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu eğitim ve doğrulama kayıpları



Şekil 4.19 5 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu hata matrisi



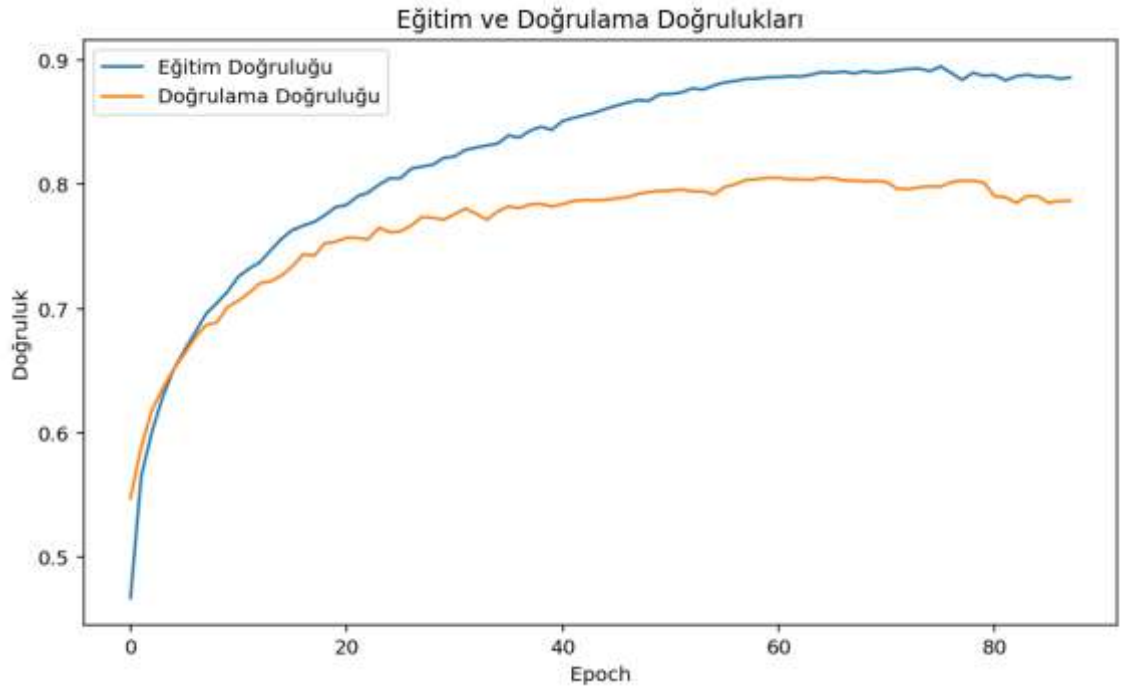
Şekil 4.20 5 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu normalize hata matrisi

Çizelge 4.5'te raporlanan sınıf bazlı metriklerde Class E (F1-Skoru = 0,9784) ve Class D (F1-Skoru = 0,8998) yine en yüksek başarıyı sergilemektedir. Class A, B ve C'nin F1-Skorları sırasıyla 0,8429, 0,8010 ve 0,8053 düzeyindedir. Örneklem hacmindeki artış, özellikle destek değeri en yüksek olan Class D'nin başarımını 10 saniyelik konfigürasyona kıyasla iyileştirmiştir.

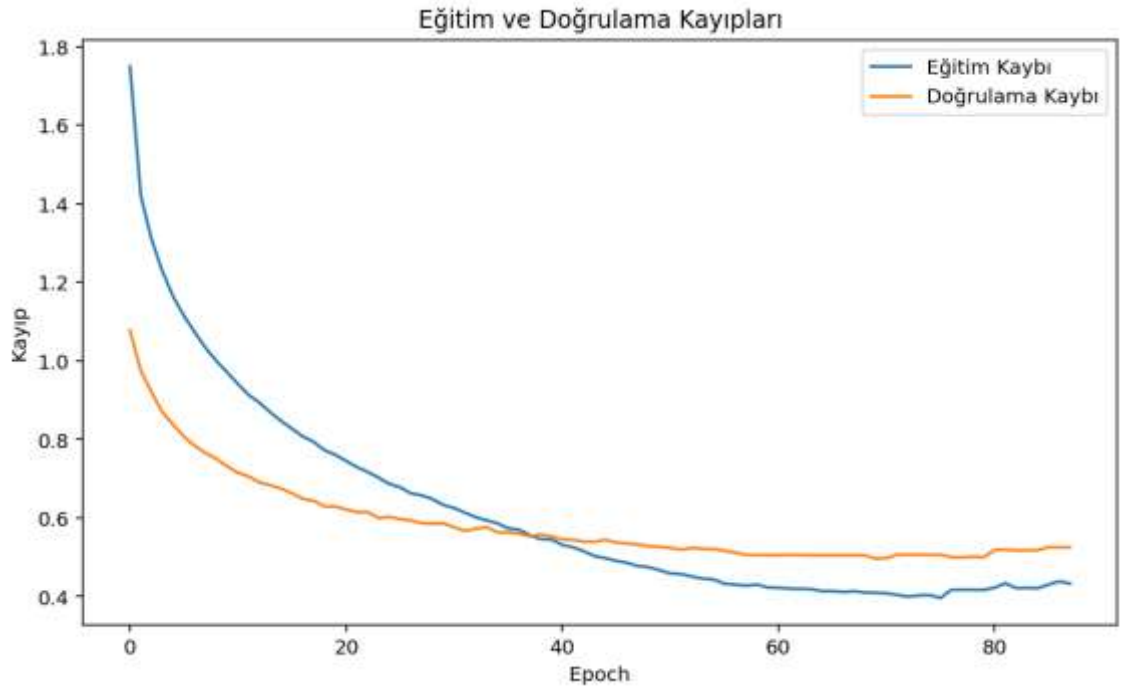
Çizelge 4.5 5 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu performans metrikleri

| Sınıflar              | Kesinlik | Duyarlılık | F1-Skoru | Destek |
|-----------------------|----------|------------|----------|--------|
| Class A               | 0,8501   | 0,8359     | 0,8429   | 658    |
| Class B               | 0,7938   | 0,8083     | 0,8010   | 600    |
| Class C               | 0,8091   | 0,8015     | 0,8053   | 534    |
| Class D               | 0,9004   | 0,8992     | 0,8998   | 1508   |
| Class E               | 0,9687   | 0,9884     | 0,9784   | 344    |
| <b>Doğruluk</b>       |          |            | 0,8669   | 3644   |
| <b>Makro Ort.</b>     | 0,8644   | 0,8667     | 0,8655   | 3644   |
| <b>Ağırlıklı Ort.</b> | 0,8668   | 0,8669     | 0,8668   | 3644   |

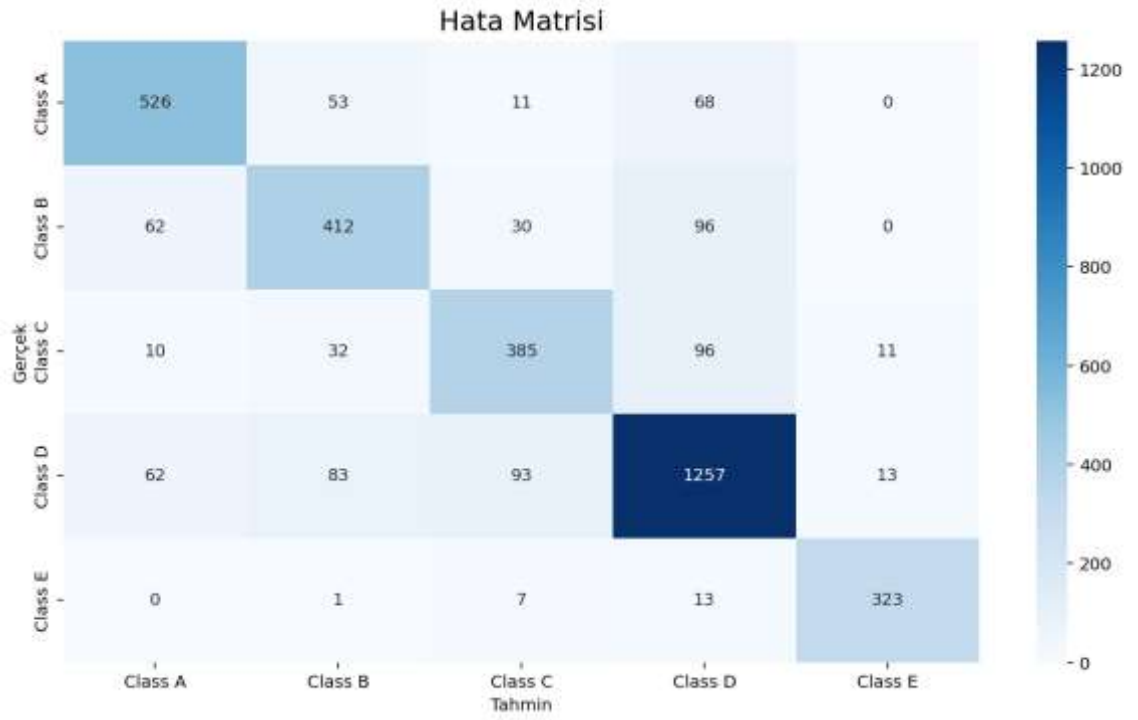
$N(3, 5^2)$  rejiminde gürültü ortalamasının 3 dB yükseltilmesiyle ortalama doğrulama doğruluğu  $\%79,67 \pm \%2,78$  düzeyine, makro F1-Skoru ise  $0,7967 \pm 0,0271$  düzeyine gerilemektedir.  $N(0, 5^2)$  rejimine kıyasla düşüş 7,02 puandır.



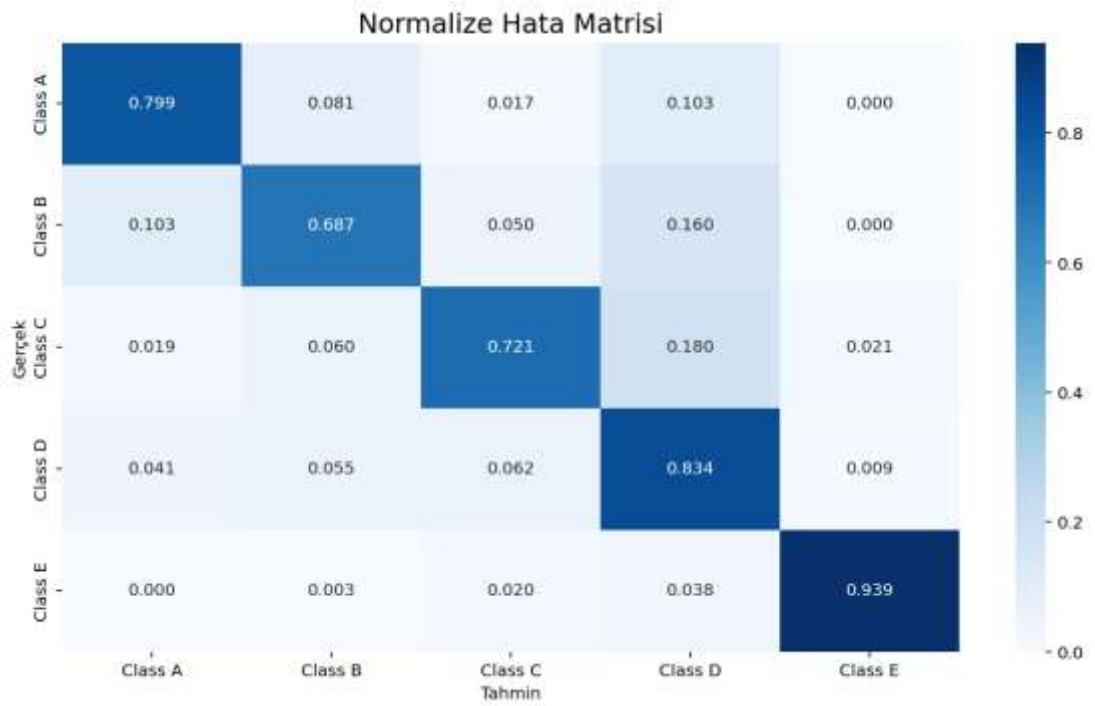
Şekil 4.21 5 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu eğitim ve doğrulama doğrulukları



Şekil 4.22 5 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu eğitim ve doğrulama kayıpları



Şekil 4.23 5 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu hata matrisi



Şekil 4.24 5 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu normalize hata matrisi

Çizelge 4.6 5 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu performans metrikleri

| Sınıflar              | Kesinlik | Duyarlılık | F1-Skoru | Destek |
|-----------------------|----------|------------|----------|--------|
| Class A               | 0,7970   | 0,7994     | 0,7982   | 658    |
| Class B               | 0,7091   | 0,6867     | 0,6977   | 600    |
| Class C               | 0,7319   | 0,7210     | 0,7264   | 534    |
| Class D               | 0,8216   | 0,8336     | 0,8275   | 1508   |
| Class E               | 0,9308   | 0,9390     | 0,9349   | 344    |
| <b>Doğruluk</b>       |          |            | 0,7967   | 3644   |
| <b>Makro Ort.</b>     | 0,7981   | 0,7959     | 0,7969   | 3644   |
| <b>Ağırlıklı Ort.</b> | 0,7958   | 0,7967     | 0,7962   | 3644   |

Gürültüsüz kontrol deneyinde gürültü artırımının başarıma katkısını anlamak amacıyla, 5 saniyelik segmentasyon aynı protokol altında herhangi bir gürültü artırımı uygulanmadan eğitilmiştir. Bu koşulda ortalama doğrulama doğruluğu  $\%82,37 \pm \%3,21$  ve makro F1-Skoru  $0,8243 \pm 0,0297$  olarak elde edilmiştir (Çizelge 4.7). Elde edilen bu değer, iki gürültü rejimiyle birlikte değerlendirildiğinde gürültü artırımının çift yönlü etkisini ortaya koymaktadır:  $N(0, 5^2)$  rejimi, gürültüsüz temele kıyasla doğruluğu 4,32 puan artırmakta ( $\%82,37 \rightarrow \%86,69$ ); buna karşılık  $N(3, 5^2)$  rejimi, doğruluğu gürültüsüz temelin 2,70 puan altına düşürmektedir ( $\%82,37 \rightarrow \%79,67$ ). Diğer bir ifadeyle, ılımlı bir gürültü artırımı (0 dB ortalama) düzenlileştirici (regularizing) bir etki sağlayarak modelin genelleme yeteneğini iyileştirmekte; aşırı yüksek şiddetli bir gürültü artırımı (3 dB ortalama) ise eğitim işaretini bozarak, gürültü artırımının hiç uygulanmadığı duruma kıyasla dahi daha düşük bir başarıma yol açmaktadır.

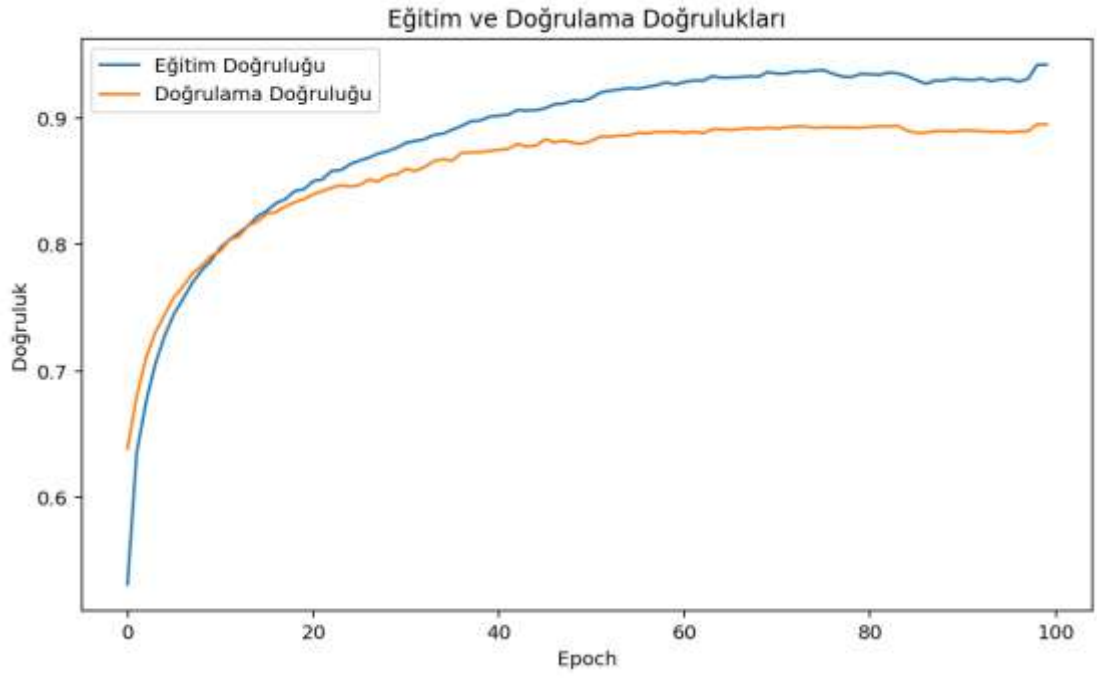
Çizelge 4.7 5 saniyelik segmentasyon deneyi (gürültü eklenmeden) sonucu performans metrikleri

| Sınıflar              | Kesinlik | Duyarlılık | F1-Skoru | Destek |
|-----------------------|----------|------------|----------|--------|
| Class A               | 0,8278   | 0,8210     | 0,8244   | 486    |
| Class B               | 0,7695   | 0,7236     | 0,7458   | 369    |
| Class C               | 0,7050   | 0,7542     | 0,7287   | 301    |
| Class D               | 0,8519   | 0,8529     | 0,8524   | 843    |
| Class E               | 0,9605   | 0,9777     | 0,9690   | 224    |
| <b>Doğruluk</b>       |          |            | 0,8237   | 2223   |
| <b>Makro Ort.</b>     | 0,8229   | 0,8259     | 0,8241   | 2223   |
| <b>Ağırlıklı Ort.</b> | 0,8240   | 0,8237     | 0,8236   | 2223   |

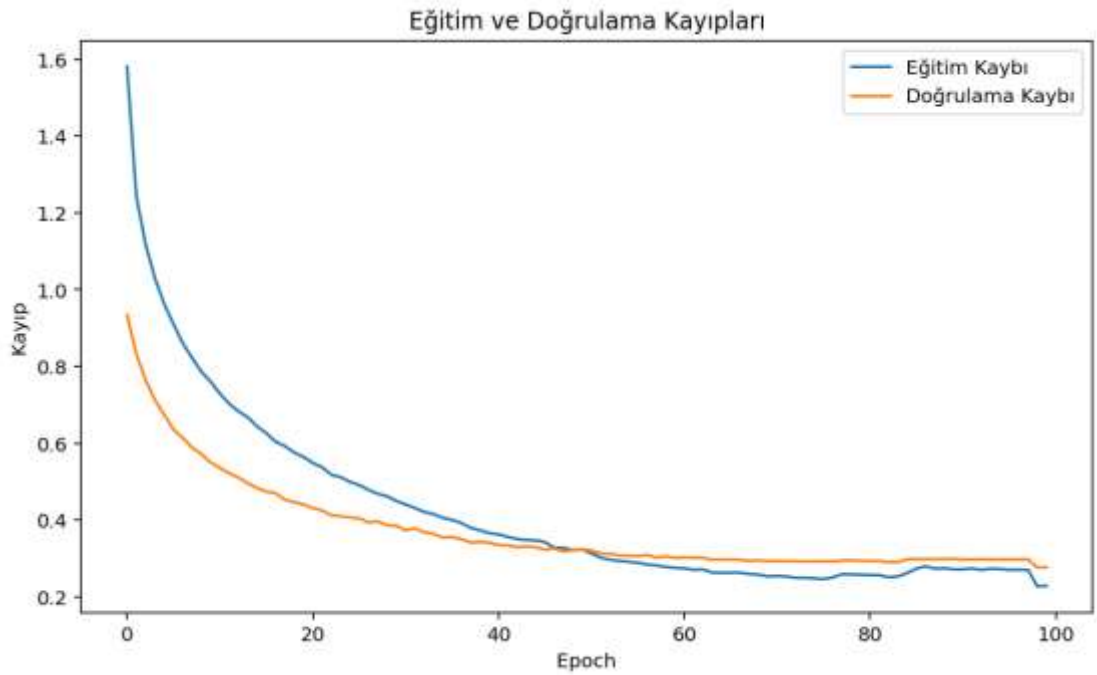
#### 4.6.3 3 saniyelik segmentasyon deneyi sonuçları

Örnek sayısının 6126'ya yükseldiği 3 saniyelik deney, her iki gürültü rejiminde de en yüksek doğrulama doğruluğunun elde edildiği deney olarak öne çıkmıştır. Bu sonuç, segment uzunluğunun akustik bağlam zenginliği ile örneklem hacmi arasında kurduğu ödünleşimin 3 saniyede en elverişli dengeye ulaştığına işaret etmektedir.

$N(0, 5^2)$  rejimi, tüm deneyler arasındaki en yüksek başarıyı vermektedir: ortalama doğrulama doğruluğu  $88,52 \pm 1,39\%$ , makro F1-Skoru ise  $0,8887 \pm 0,0131$  olarak elde edilmiştir. Düşük standart sapma ( $\pm 1,39\%$ ), yüksek doğrulukla birlikte katlamalar arası yüksek bir kararlılığa işaret etmektedir. Şekil 4.25 ve Şekil 4.26'daki öğrenme eğrileri, doğrulama doğruluğunun 0,88–0,90 bandında dengelendiğini ve eğitim–doğrulama açıklığının diğer segment uzunluklarına kıyasla daraldığını ortaya koymaktadır. Bu daralma, artan örneklem hacminin ezberleme eğilimini sınırladığına işaret etmektedir.

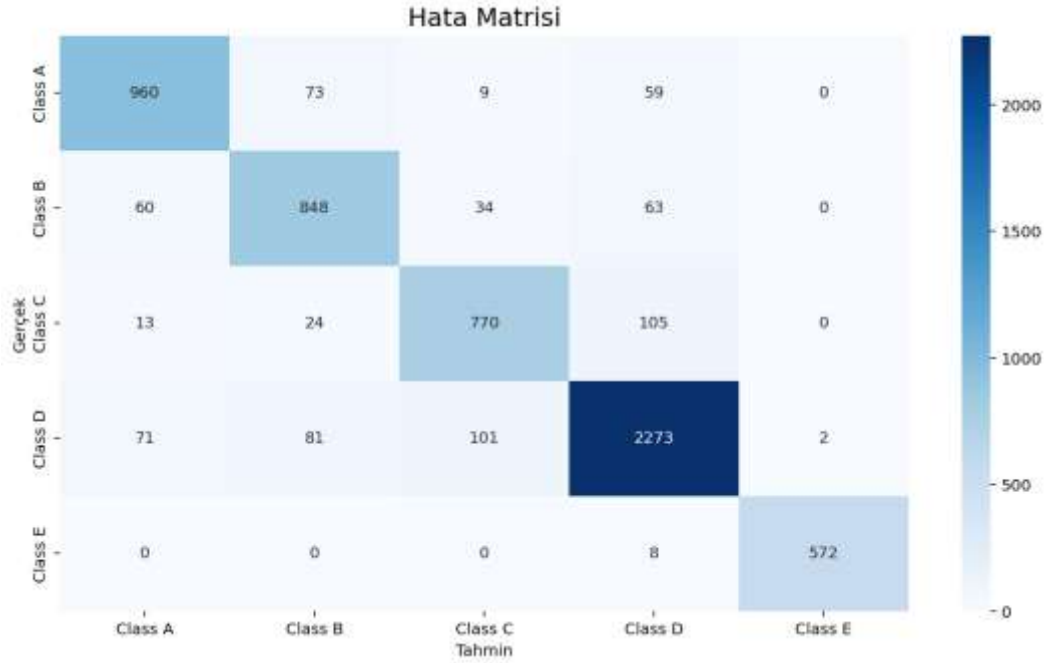


Şekil 4.25 3 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu eğitim ve doğrulama doğrulukları

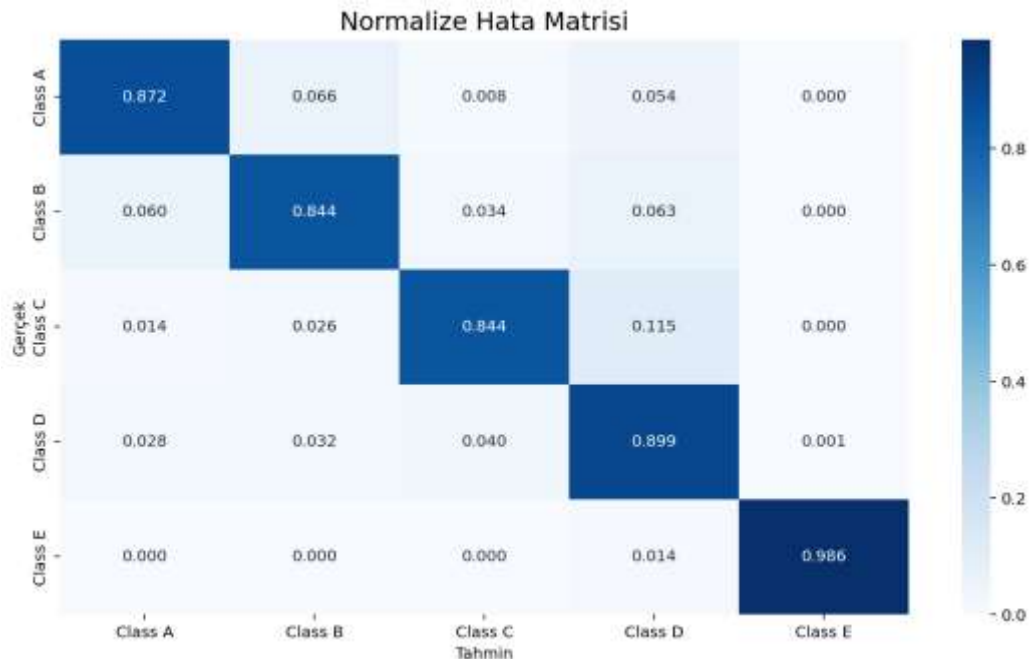


Şekil 4.26 3 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu eğitim ve doğrulama kayıpları

Şekil 4.25 ve Şekil 4.26'da görüldüğü üzere, eğitim ve doğrulama eğrilerinin yakınsama davranışı dört deney arasında en stabil profili sergilemektedir. Doğrulama kaybındaki dalgalanmalar son derece sınırlı kalmıştır.



Şekil 4.27 3 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu hata matrisi



Şekil 4.28 3 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu normalize hata matrisi

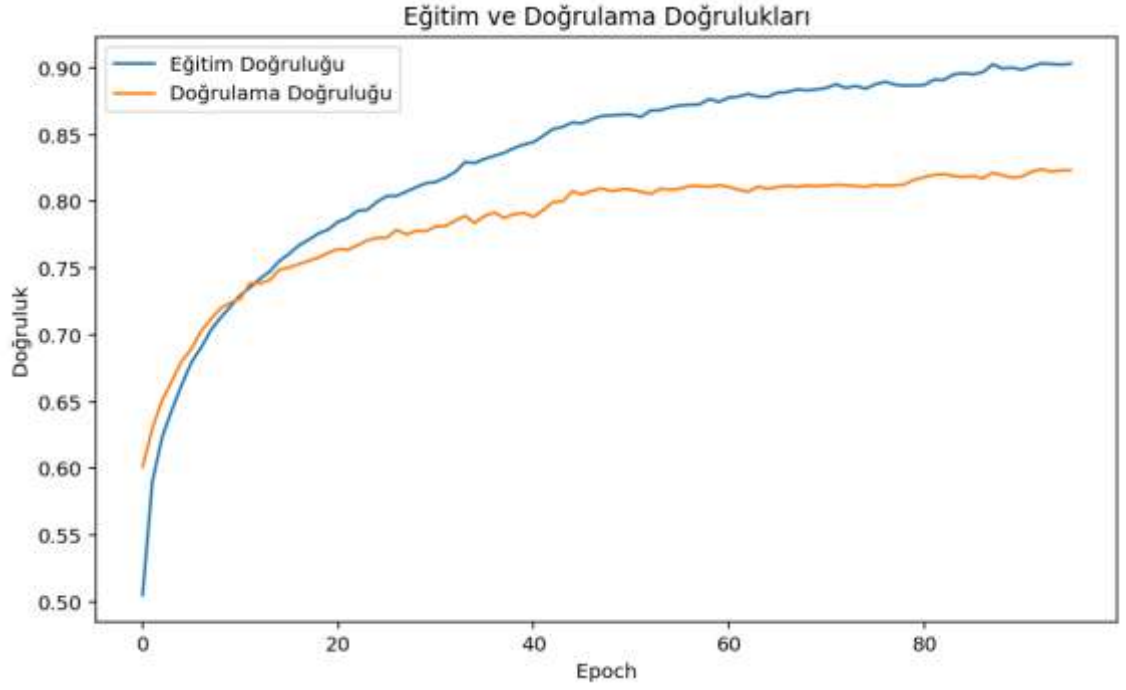
Hata matrisinin sol üstten sağ alta uzanan köşegenindeki değerlerin artması, 3 saniyelik segmentasyonun gemi kaynaklı periyodik bileşenleri (pervane vuruşu, motor harmonikleri gibi) tek bir spektrogramda yeterli zaman bağlamıyla temsil etmesiyle açıklanabilir.

Çizelge 4.8'de görüldüğü üzere, bu konfigürasyonda tüm sınıflar en yüksek F1-Skorlarına ulaşmaktadır. Class E 0,9913 ile en yüksek ayırt edilebilirliği sergilerken, en zorlu sınıflar olan Class B ve Class C dahi sırasıyla 0,8351 ve 0,8434 düzeyine yükselmektedir. Bu durum, 3 saniyelik bağlamın; düşük-orta frekans bandında örtüşen küçük ve orta ölçekli araçların ayırt edici örüntülerini yakalamaya yeterli süreyi sağlaması ve aynı zamanda yeterli sayıda eğitim örneği üretmesiyle açıklanabilir.

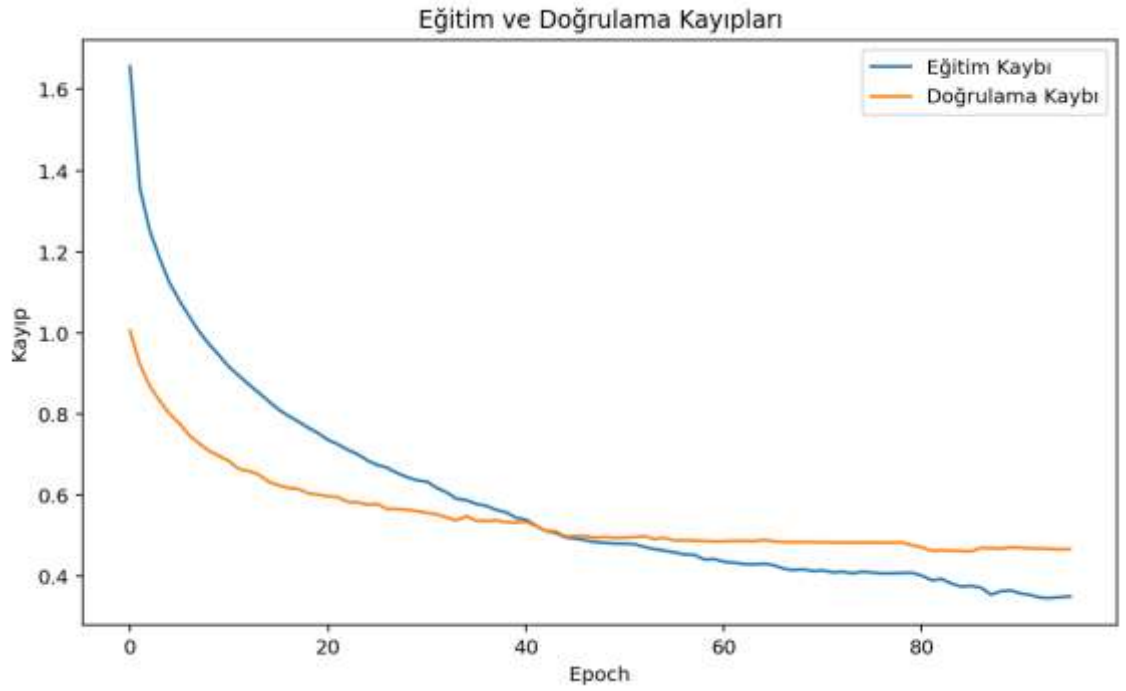
Çizelge 4.8 3 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu performans metrikleri

| Sınıflar              | Kesinlik | Duyarlılık | F1-Skoru | Destek |
|-----------------------|----------|------------|----------|--------|
| Class A               | 0,8696   | 0,8719     | 0,8707   | 1101   |
| Class B               | 0,8265   | 0,8438     | 0,8351   | 1005   |
| Class C               | 0,8425   | 0,8443     | 0,8434   | 912    |
| Class D               | 0,9063   | 0,8991     | 0,9027   | 2528   |
| Class E               | 0,9965   | 0,9862     | 0,9913   | 580    |
| <b>Doğruluk</b>       |          |            | 0,8852   | 6126   |
| <b>Makro Ort.</b>     | 0,8883   | 0,8891     | 0,8886   | 6126   |
| <b>Ağırlıklı Ort.</b> | 0,8856   | 0,8852     | 0,8854   | 6126   |

3 saniyelik konfigürasyon,  $N(3, 5^2)$  rejimi içinde de en yüksek doğruluğu vermektedir: ortalama doğrulama doğruluğu  $\%80,51 \pm \%1,77$ , makro F1-Skoru ise  $0,8084 \pm 0,0199$  olarak elde edilmiştir. Bu sonuç, segment uzunluğunun optimum değerinin gürültü rejiminden bağımsız olarak 3 saniyede kaldığını; yani segment uzunluğu ile gürültü şiddetinin başarımlar üzerindeki etkilerinin büyük ölçüde ayrışabilir (birbirinden bağımsız) olduğunu göstermesi bakımından önemlidir.



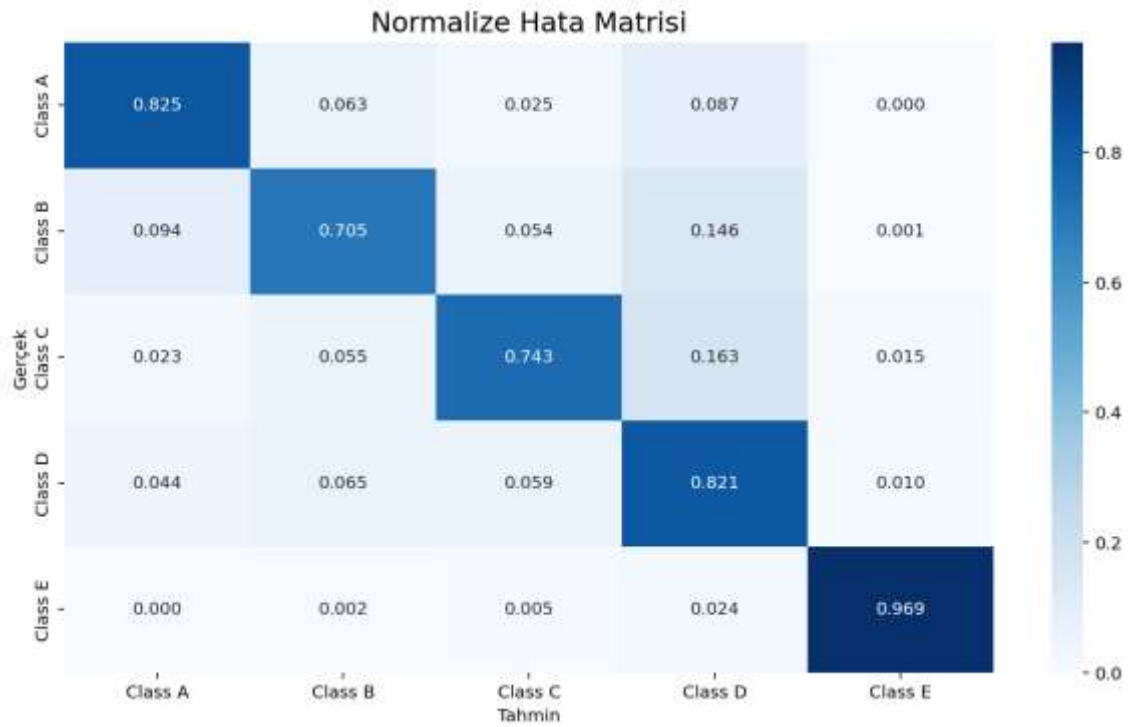
Şekil 4.29 3 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu eğitim ve doğrulama doğrulukları



Şekil 4.30 3 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu eğitim ve doğrulama kayıpları



Şekil 4.31 3 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu hata matrisi



Şekil 4.32 3 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu normalize hata matrisi

Çizelge 4.9'da, gürültü altında dahi Class E'nin ( $F1 = 0,9501$ ) yüksek başarımını koruduğu; buna karşılık Class B'nin ( $F1 = 0,7094$ ) en fazla gerileyen sınıf olduğu görülmektedir. Bu örüntü, önceki segment uzunluklarında gözlenen sınıf duyarlılık sıralamasıyla tutarlıdır.

Çizelge 4.9 3 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu performans metrikleri

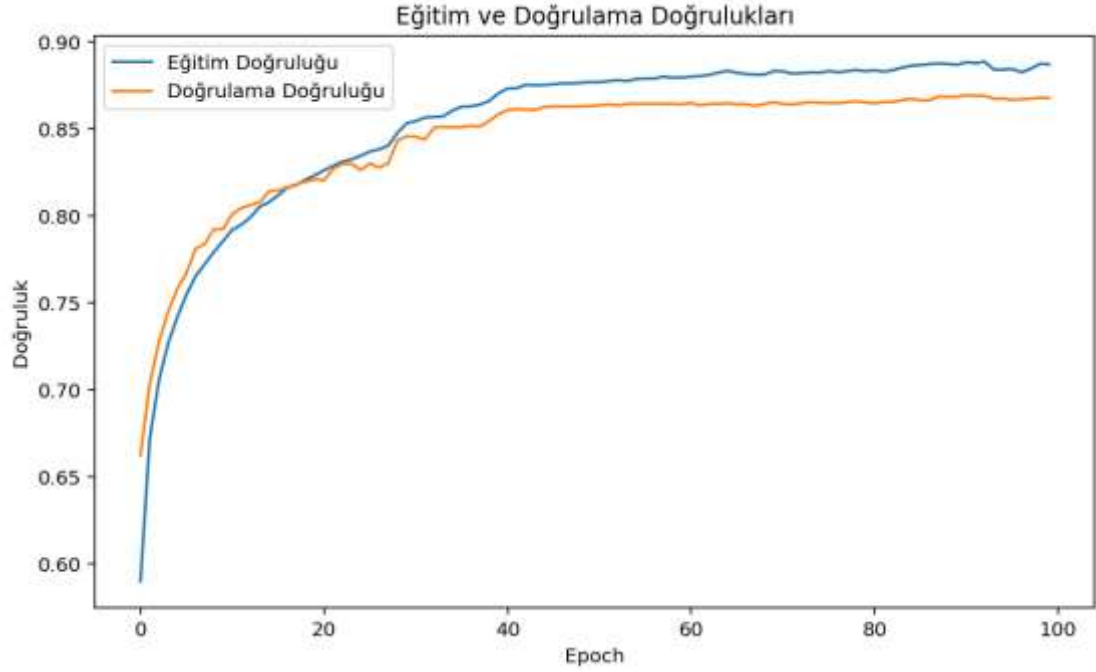
| Sınıflar              | Kesinlik | Duyarlılık | F1-Skoru | Destek |
|-----------------------|----------|------------|----------|--------|
| Class A               | 0,8000   | 0,8247     | 0,8122   | 1101   |
| Class B               | 0,7133   | 0,7055     | 0,7094   | 1005   |
| Class C               | 0,7426   | 0,7434     | 0,7430   | 912    |
| Class D               | 0,8364   | 0,8208     | 0,8285   | 2528   |
| Class E               | 0,9320   | 0,9690     | 0,9501   | 580    |
| <b>Doğruluk</b>       |          |            | 0,8051   | 6126   |
| <b>Makro Ort.</b>     | 0,8048   | 0,8127     | 0,8086   | 6126   |
| <b>Ağırlıklı Ort.</b> | 0,8047   | 0,8051     | 0,8048   | 6126   |

#### 4.6.4 1 saniyelik segmentasyon deneyi sonuçları

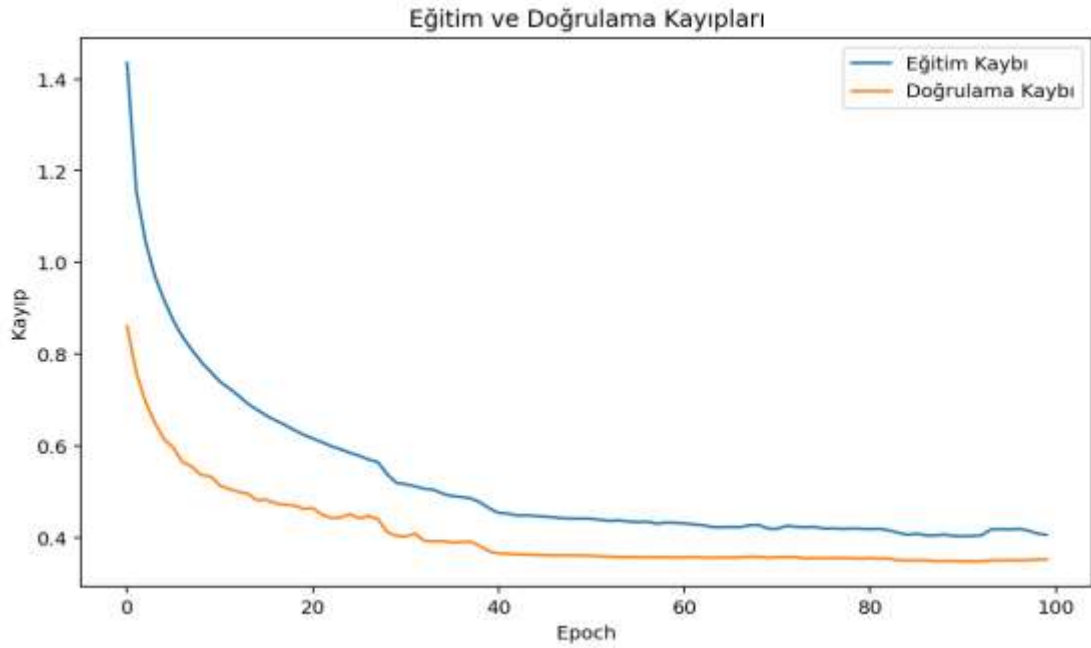
1 saniyelik segmentasyon deneyi, örnek sayısını 18524 spektrograma çıkararak en büyük örneklem hacmini üretmiştir. Ancak, segment uzunluğunun bu denli kısılması, modelin akustik imzayı yakaladığı zaman bağlamını da daraltmıştır.

$N(0, 5^2)$  rejiminde ortalama doğrulama doğruluğu  $\%86,02 \pm \%0,72$ , makro F1-Skoru ise  $0,8636 \pm 0,0083$  olarak elde edilmiştir. Bu konfigürasyon, tüm deneyler arasındaki en düşük standart sapmayı ( $\pm \%0,72$ ) vermektedir; bu durum, en yüksek katlamalar arası kararlılığın büyük örneklem hacminden kaynaklandığını desteklemektedir. Bununla birlikte ortalama doğruluk, 3 saniyelik konfigürasyonun ( $\%88,52$ ) 2,50 puan altında kalmaktadır. Standart sapmanın düşük bir değere ( $\pm \%0,72$ ) inmesine rağmen ortalama doğrulama doğruluğunun 3 saniyelik segmentasyon deneyine kıyasla 2,50 puan gerilemesi, daha kısa segment pencerelerinin sınıf ayırıcı bağlam bilgisini yeterince temsil edemediğini göstermektedir. Şekil 4.33 ve Şekil 4.34'teki öğrenme eğrileri, eğitim

ve doğrulama eğrilerinin birbirine en yakın seyrettiği; yani ezberleme eğiliminin en aza indiği durumu göstermektedir.



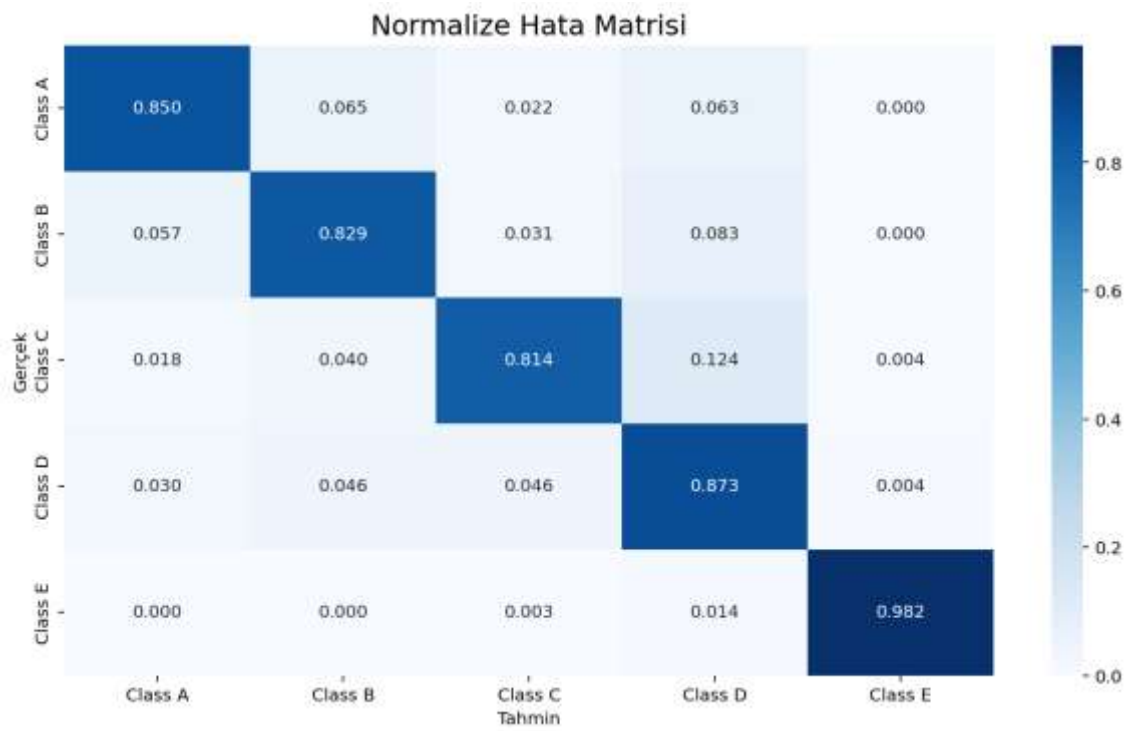
Şekil 4.33 1 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu eğitim ve doğrulama doğrulukları



Şekil 4.34 1 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu eğitim ve doğrulama kayıpları



Şekil 4.35 1 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu hata matrisi



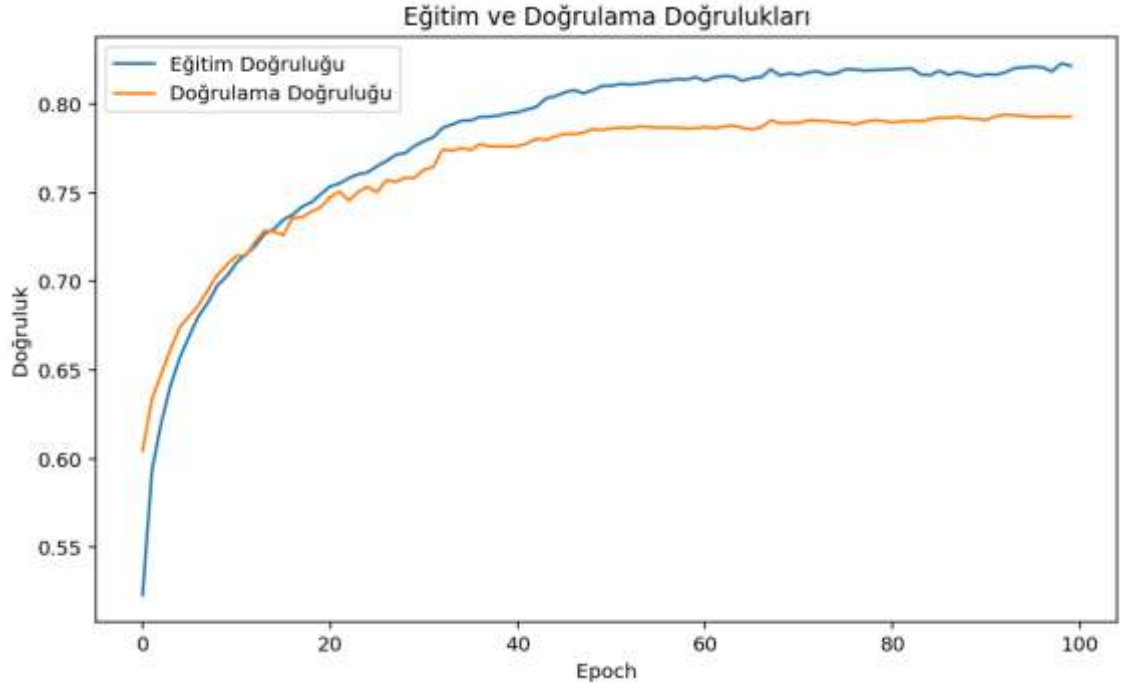
Şekil 4.36 1 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu normalize hata matrisi

Çizelge 4.10'da raporlanan sınıf bazlı metriklerde, 3 saniyelik segmentasyon deneyine göre ( $N(0, 5^2)$  rejimi) performans düşüşü en belirgin biçimde gemi sınıflarında (A, B, C, D) görülmektedir; Class A F1-Skoru 0,8707'den 0,8497'ye, Class B 0,8351'den 0,8198'e, Class C 0,8434'ten 0,7955'e ve Class D 0,9027'den 0,8787'ye gerilemiştir. Class E için ise bu gerileme oldukça sınırlı kalmıştır ( $0,9913 \rightarrow 0,9744$ ). Bu farklılaşmış davranış, segment uzunluğunun düşürülmesinin tüm sınıflar üzerinde aynı etkiyi oluşturmadığını ortaya koymaktadır.

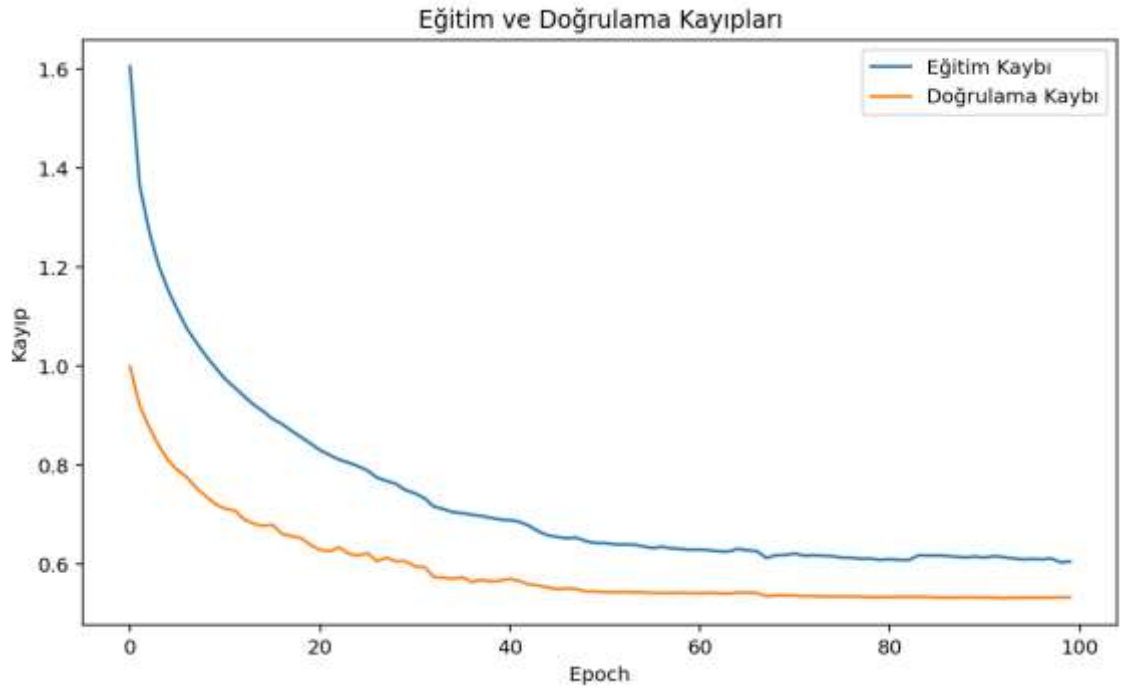
Çizelge 4.10 1 saniyelik segmentasyon deneyi ( $N(0, 5^2)$  rejimi) sonucu performans metrikleri

| Sınıflar              | Kesinlik | Duyarlılık | F1-Skoru | Destek |
|-----------------------|----------|------------|----------|--------|
| Class A               | 0,8467   | 0,8528     | 0,8497   | 3322   |
| Class B               | 0,8084   | 0,8315     | 0,8198   | 3050   |
| Class C               | 0,7917   | 0,7994     | 0,7955   | 2767   |
| Class D               | 0,8875   | 0,8701     | 0,8787   | 7635   |
| Class E               | 0,9711   | 0,9777     | 0,9744   | 1750   |
| <b>Doğruluk</b>       |          |            | 0,8602   | 18524  |
| <b>Makro Ort.</b>     | 0,8611   | 0,8663     | 0,8636   | 18524  |
| <b>Ağırlıklı Ort.</b> | 0,8607   | 0,8602     | 0,8604   | 18524  |

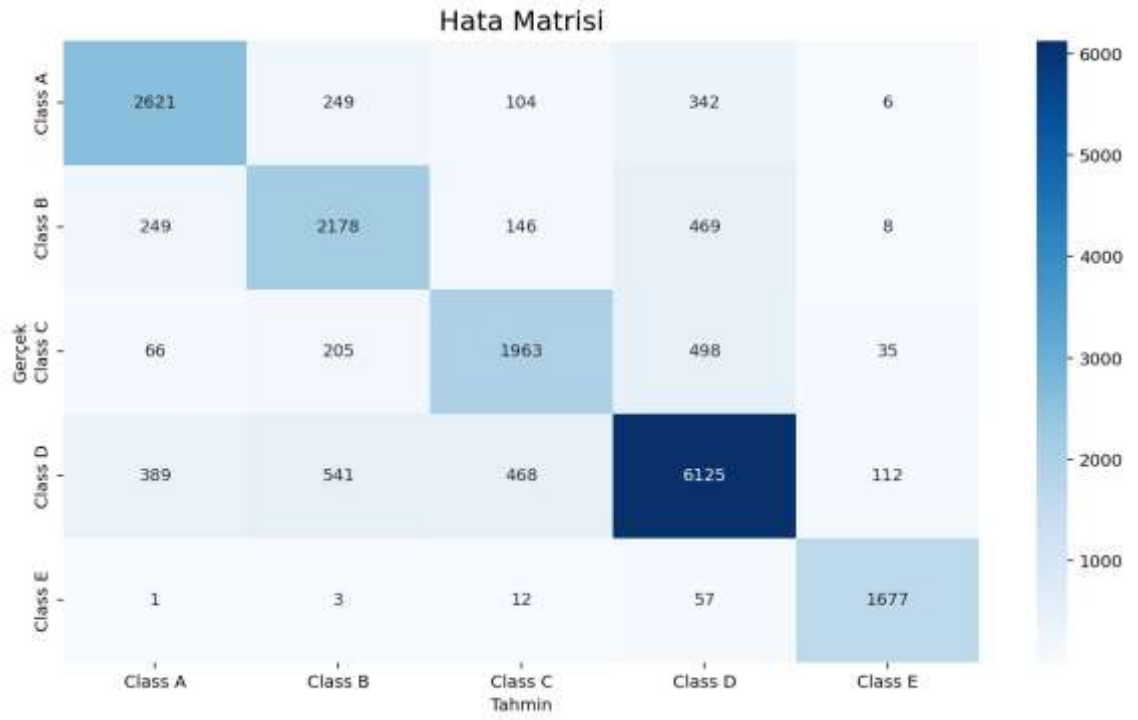
$N(3, 5^2)$  rejiminde gürültü ortalamasının yükseltilmesiyle ortalama doğrulama doğruluğu  $\%78,62 \pm \%1,52$ , makro F1-Skoru  $0,7903 \pm 0,0149$  düzeyine gerilemektedir.



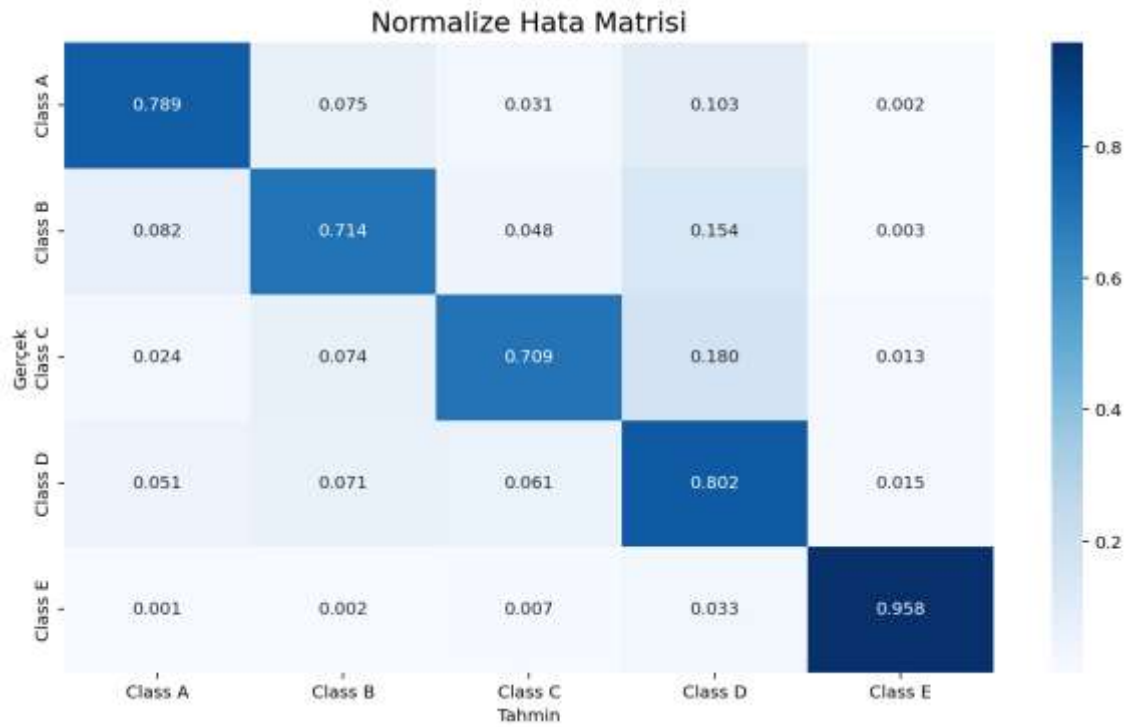
Şekil 4.37 1 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu eğitim ve doğrulama doğrulukları



Şekil 4.38 1 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu eğitim ve doğrulama kayıpları



Şekil 4.39 1 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu hata matrisi



Şekil 4.40 1 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu normalize hata matrisi

Çizelge 4.11'de, gürültü altında Class B ( $F1 = 0,6996$ ) ve Class C ( $F1 = 0,7190$ ) yeniden en düşük başarımları sergilerken, Class E ( $F1 = 0,9348$ ) dayanıklılığını korumaktadır.

Çizelge 4.11 1 saniyelik segmentasyon deneyi ( $N(3, 5^2)$  rejimi) sonucu performans metrikleri

| Sınıflar              | Kesinlik | Duyarlılık | F1-Skoru | Destek |
|-----------------------|----------|------------|----------|--------|
| Class A               | 0,7880   | 0,7890     | 0,7885   | 3322   |
| Class B               | 0,6858   | 0,7141     | 0,6996   | 3050   |
| Class C               | 0,7289   | 0,7094     | 0,7190   | 2767   |
| Class D               | 0,8176   | 0,8022     | 0,8099   | 7635   |
| Class E               | 0,9124   | 0,9583     | 0,9348   | 1750   |
| <b>Doğruluk</b>       |          |            | 0,7862   | 18524  |
| <b>Makro Ort.</b>     | 0,7866   | 0,7946     | 0,7904   | 18524  |
| <b>Ağırlıklı Ort.</b> | 0,7863   | 0,7862     | 0,7861   | 18524  |

#### 4.6.5 Segment uzunluğu ve gürültü rejiminin model performansına etkileri üzerine karşılaştırmalı analiz

Bu çalışmada benimsenen iki değişkenli faktöriyel tasarım, segment uzunluğunun (10, 5, 3 ve 1 saniye) ve eğitim setine uygulanan Gauss gürültüsü rejiminin ( $N(0, 5^2)$  ve  $N(3, 5^2)$ ) model başarımları üzerindeki etkilerinin; ön işleme adımları, ağ mimarisi, eğitim hiperparametreleri ve tabakalı 10 katlı çapraz doğrulama protokolü sabit tutularak birlikte incelenmesine olanak tanımaktadır. Buna ek olarak yürütülen gürültüsüz kontrol deneyi, gürültü artırımının başarıma net katkısının anlaşılmasını sağlamaktadır. Bu alt bölümde, önceki dört alt bölümde ayrı ayrı raporlanan bulgular Çizelge 4.2 ekseninde bir araya getirilerek iki tasarım değişkeninin bağımsız ana etkileri, aralarındaki olası etkileşim ve gürültü artırımının çift yönlü niteliği karşılaştırmalı biçimde değerlendirilmektedir.

Segment uzunluğunun ana etkisi incelendiğinde, doğrulama doğruluğu ile segment uzunluğu arasındaki ilişkinin tekdüze (monoton) olmadığı ve tepe noktası 3 saniyede

bulunan bir örüntü sergilediği görülmektedir.  $N(0, 5^2)$  rejiminde ortalama doğrulama doğruluğu, segment uzunluğu 10 saniyeden 3 saniyeye indikçe  $\%85,92 \rightarrow \%86,69 \rightarrow \%88,52$  seyriyle yükselmekte; pencerenin 1 saniyeye daraltılmasıyla  $\%86,02$  düzeyine gerilemektedir.  $N(3, 5^2)$  rejiminde de aynı örüntü korunmakta, doğruluk  $\%77,96 \rightarrow \%79,67 \rightarrow \%80,51$  seyriyle yine 3 saniyede tepe yapmakta ve 1 saniyede  $\%78,62$  düzeyine düşmektedir. Optimum segment uzunluğunun her iki gürültü rejiminde de değişmeksizin 3 saniyede kalması, segment uzunluğunun etkisinin gürültü rejiminden büyük ölçüde bağımsız (ayrışabilir) olduğuna işaret etmektedir. 3 saniyelik konfigürasyonun birbirinden bağımsız iki rejimde birden tepe noktasını oluşturması, bu sonucun tek bir ikili karşılaştırmaya değil tutarlı bir örüntüye dayandığını göstermesi bakımından önemlidir.

Uzun segmentler (10 saniye) zengin bir zaman bağlamı sunmakla birlikte 90 ham kayıttan yalnızca 1797 örnek üretebilmekte; bu sınırlı örneklem hacmi, katlamalar arasında en yüksek standart sapmaya ( $N(0, 5^2)$  rejiminde  $\pm\%3,14$ ) yol açarak yüksek varyansla cezalandırılmaktadır. Segment uzunluğu kısaldıkça örneklem hacmi büyümekte (5 saniyede 3644, 3 saniyede 6126 ve 1 saniyede 18524 spektrograma ulaşmakta) ve katlamalar arası standart sapma her iki rejimde de tekdüze biçimde azalmaktadır;  $N(0, 5^2)$  rejiminde bu azalış  $\pm\%3,14 \rightarrow \pm\%1,47 \rightarrow \pm\%1,39 \rightarrow \pm\%0,72$  biçiminde gerçekleşmektedir. En düşük standart sapmanın gözlemlendiği 1 saniyelik konfigürasyon ( $\pm\%0,72$ ), en yüksek ortalama doğruluğu vermemektedir. Bu durum, örneklem hacminin sağladığı istatistiksel kararlılığın (düşük varyansın) tek başına yüksek sınıflandırma başarımının güvencesi olmadığını; kısa pencerelerin sınıf ayırıcı bağlamı yeterince temsil edememesi nedeniyle modelin yüksek bir yanlılıkla (bias) sınırlandığını ortaya koymaktadır.

1 saniyelik konfigürasyonda gözlenen başarımların kaybolmasının kaynağı, model mimarisi veya eğitim protokolü değil, doğrudan girdinin temsil ettiği fiziksel zaman penceresidir. Gemi kaynaklı akustik imzalar; pervane dönüş döngüsü, makine harmonikleri ve baskın tonal bileşenler gibi belirli bir asgari süre boyunca gözlemlendiğinde ayırt edici hâle gelen periyodik yapılar üzerinden tanımlanmaktadır. 1 saniyelik bir pencere bu periyodik yapıları tek bir spektrogramda yeterli bağlamla temsil edemediğinden, örnek sayısı artsa

dahi her bir örneğin taşıdığı sınıf ayırıcı bilgi miktarı azalmaktadır. Buna karşılık 3 saniyelik pencere, söz konusu periyodik bileşenleri temsil etmeye yetecek zaman bağlamını korurken aynı zamanda modelin genelleme yeteneği için gerekli örnek çeşitliliğini de sağlayarak iki uç arasındaki en elverişli dengeyi kurmaktadır.

Gürültü rejiminin ana etkisi incelendiğinde,  $N(3, 5^2)$  rejiminin tüm segment uzunluklarında  $N(0, 5^2)$  rejiminin belirgin biçimde altında bir başarımla verildiği görülmektedir. İki rejim arasındaki ortalama doğruluk farkı 10, 5, 3 ve 1 saniyelik konfigürasyonlarda sırasıyla 7,96; 7,02; 8,01 ve 7,40 puan olup yaklaşık 7-8 puanlık dar bir bantta seyretmektedir. Bu farkın segment uzunluğundan görece bağımsız ve yaklaşık sabit kalması, iki tasarım değişkeninin başarımla üzerindeki etkilerinin belirgin bir etkileşim içermeksizin büyük ölçüde toplamsal (additive) bir karakter taşıdığını desteklemektedir.  $N(3, 5^2)$  rejimi, eğitim örneklerine ortalamada 3 dB daha yüksek gürültü ekleyerek etkin sinyal-gürültü oranını düşürmekte ve sınıf ayırıcı bilgiyi kısmen örtük olarak öğrenme problemini tüm segment uzunluklarında benzer ölçüde zorlaştırmaktadır.

Gürültü artırımının başarıma katkısının yönü ve büyüklüğü, gürültüsüz kontrol deneyi referans alındığında daha ayrıntılı biçimde çözümlenebilmektedir. 5 saniyelik segmentasyon için herhangi bir gürültü artırımı uygulanmadığında ortalama doğrulama doğruluğu %82,37 düzeyinde kalmaktadır. Aynı segment uzunluğunda  $N(0, 5^2)$  rejimi bu temel değeri 4,32 puan yükseltirken (%82,37  $\rightarrow$  %86,69),  $N(3, 5^2)$  rejimi başarımla temel değerin 2,70 puan altına düşürmektedir (%82,37  $\rightarrow$  %79,67). Bu karşılaştırma, gürültü artırımının tek yönlü ve monoton biçimde yararlı bir işlem olmadığını; etkisinin, eklenen gürültünün dağılım ortalamasına, dolayısıyla ortalama etkin sinyal-gürültü oranına bağlı olarak yön değiştirdiğini göstermektedir. Ortalaması 0 dB olan ılımlı bir gürültü artırımı, modeli spektrogramlardaki gürültü örüntülerine karşı dayanıklı kılan düzenlileştirici (regularizing) bir etki sağlarken; ortalaması 3 dB olan aşırı şiddetli bir artırım, eğitim işaretini bozarak gürültünün hiç eklenmediği duruma kıyasla dahi daha düşük bir başarıma yol açmaktadır.

Sınıf bazlı başarımların örüntüleri de segment uzunluğundan ve gürültü rejiminden bağımsız, yapısal olarak kararlı bir sıralama ortaya koymaktadır. Doğal ortam gürültüsünü temsil eden ve tonal hat ya da sistematik harmonik bileşen içermeyen geniş bantlı yapısıyla Class E, tüm konfigürasyonlarda 0,93'ün üzerindeki F1-Skoru ile hem en yüksek ayırt edilebilirliği hem de en yüksek gürültü dayanıklılığını sergilemiştir. Buna karşılık spektral içerikleri özellikle 0–4 kHz düşük-orta frekans bandında örtüşen Class B ve Class C, tüm segment uzunluklarında ve her iki rejimde en düşük başarımları ve gürültüye karşı en yüksek duyarlılığı göstermiştir. Bu sıralamanın deney koşullarından bağımsız biçimde korunması, söz konusu örüntünün deneysel bir yan üründen değil, sınıfların içsel akustik ayırt edilebilirlik derecelerinden kaynaklandığına işaret etmektedir.

Bu bulgular bütün olarak değerlendirildiğinde, çalışma iki temel çıkarımı desteklemektedir. Birincisi, segment uzunluğu yalnızca örneklem hacmini belirleyen bir veri artırma parametresi değil; modele aktarılan akustik bağlamın fiziksel anlamlılığını gözetten bir optimizasyon parametresidir. En büyük örneklem hacmini üreten 1 saniyelik konfigürasyon, bu yaklaşık on katlık örnek artışına rağmen 3 saniyelik optimumun 2,50 puan altında kalmış; ortalama başarımları 10 saniyelik konfigürasyonunkine (%85,92) neredeyse özdeş bir düzeyde (%86,02) seyretmiştir. Bu sonuç, “ne kadar çok örnek, o kadar iyi” sezgisinin, sinyallerin ayırt ediciliğinin fiziksel zaman ölçeklerine bağlı olduğu su altı akustiği gibi alanlarda geçerliliğini yitirdiğini göstermektedir. İkincisi, gürültü artırımı dikkatle kalibre edilmesi gereken çift yönlü bir araçtır; ılımlı bir artırımı genellemeyi iyileştiren bir düzenleştirici işlevi görürken, aşırı şiddetli bir artırımı bozarak başarımları gürültüsüz temelin dahi altına çekmektedir. ShipsEar veri seti ve EfficientNetB1 mimarisi özelinde, en yüksek ortalama doğrulama doğruluğunu (%88,52) ve en yüksek makro F1-Skorunu (0,8887) veren, aynı zamanda uzun segmentlere kıyasla belirgin biçimde düşük bir standart sapma ( $\pm 1,39$ ) sunan 3 saniyelik segmentasyon ile  $N(0, 5^2)$  gürültü rejiminin birlikte kullanımı, bu çalışma kapsamındaki optimal konfigürasyonu oluşturmaktadır. Elde edilen bulgular, benzer veri setleri üzerinde yürütülecek gelecek araştırmalarda hem segment uzunluğunun hem de gürültü artırımı şiddetinin, girdinin fiziksel anlamlılığını ve etkin sinyal-gürültü oranını koruyacak birer optimizasyon parametresi olarak ele alınması gerektiğini güçlü biçimde desteklemektedir.

## 5. SONUÇ VE ÖNERİLER

Bu tez çalışmasında, su altı akustik hedef sınıflandırma alanında geleneksel sonar operatörü uzmanlığına dayanan klasik yaklaşımın taşıdığı sınırlılıklar ele alınmıştır. Mevcut sonar sistemlerinde analiz ve sınıflandırma süreçleri büyük ölçüde operatörlerin bilgi ve tecrübesine dayanmaktadır. DSH ve Mavi Vatan doktrini çerçevesinde Türkiye'nin deniz yetki alanlarının korunması ve stratejik çıkarların güvence altına alınması, yenilikçi çözümlerin geliştirilmesini zorunlu kılmaktadır. Operatörün sürekli izleme zorunluluğu, zamana bağlı yorgunluk, psikofizyolojik durumlardan kaynaklanan kararsızlıklar ve son derece maliyetli/uzun süreli eğitim süreçleri, klasik yaklaşımların operasyonel verimliliğini kısıtlayan temel unsurlardır. Bu durum, özellikle asimetrik harp koşullarında hızlı ve doğru karar alma süreçlerini zorlaştırmakta, insan kaynaklı hata riskini artırmaktadır. Bu bağlamda, yapay zekâ destekli otomatik sınıflandırma sistemleri, operasyonel etkinliği artırma ve insan faktöründen kaynaklanan sınırlılıkları giderme potansiyeli sunmaktadır.

Su altı ortamının karmaşık doğası, arka plan gürültüsü, frekansa bağlı soğurma, çoklu yol etkileri ve Doppler etkisi gibi faktörler sinyal işlemeyi zorlaştırmaktadır. Ayrıca, yüksek kaliteli, etiketli askeri akustik veriye erişimdeki kısıtlılıklar, su altı akustik hedef tanıma alanındaki çalışmaları zorlaştıran önemli bir engeldir.

Bu çalışma, ShipsEar veri kümesi üzerinde, önceden ImageNet ile eğitilmiş EfficientNetB1 mimarisi kullanılarak transfer öğrenme tabanlı bir CNN modelinin uygulanmasına dayanmaktadır. Araştırmanın metodolojik özgünlüğü, model başarımını etkileyen iki tasarım değişkeninin, yani ses sinyallerinin bölütlendiği segment uzunluğu ile eğitim setine uygulanan Gauss gürültüsü eklenmesi rejiminin, birbirinden yalıtılmış biçimde değil birlikte ele alındığı faktöriyel bir deney tasarımından ileri gelmektedir. Bu tasarımda segment uzunluğu 10, 5, 3 ve 1 saniye olmak üzere dört düzeyde; gürültü rejimi ise  $N(0, 5^2)$  ve  $N(3, 5^2)$  olmak üzere iki düzeyde ele alınmış ve böylece sekiz ana deney oluşturulmuştur. Ön işleme adımları, ağ mimarisi, eğitim hiperparametreleri ve tabakalı 10 katlı çapraz doğrulama protokolü tüm deneylerde sabit tutularak, gözlenen başarımların farklılıklarının yalnızca bu iki değişkenle ilişkilendirilebilmesi sağlanmıştır. Sekiz ana

deneye ek olarak, gürültü artırımının başarıma net katkısını yalıtılabilmek amacıyla 5 saniyelik segmentasyon için gürültünün hiç eklenmediği bir kontrol deneyi yürütülmüştür.

Geliştirilen ön işleme zinciri, çalışmanın yeniden üretilebilirliği açısından somut bir metodolojik şablon sunmaktadır. Bu kapsamda ses kayıtları 16 kHz'e yeniden örneklenmiş, kayıt koşulu ve mesafe farklılıklarından kaynaklanan enerji değişkenliğini gidermek amacıyla RMS normalizasyonu uygulanmış ve desibel cinsinden önceden belirlenen sabit global minimum ve maksimum güç değerleri ortak bir renk skalası olarak kullanılmıştır. Bu sabit ölçek, tüm spektrogramların aynı dinamik aralıkta görselleştirilmesini sağlayarak sınıflar arası ve sınıf içi karşılaştırılabilirliği yapısal olarak korumuştur. Gürültü artırımı yalnızca eğitim setine uygulanmış; doğrulama ve test kümeleri ham halleriyle korunarak test başarımının modelin gerçek sınıflandırma yeteneğini yansıtmayı ve eklenen gürültü dağılımına aşırı uyum oluşması engellenmiştir.

Veri artırma stratejisinin seçimi, sinyalin fiziksel doğasıyla uyum ilkesine dayandırılmıştır. Görüntü sınıflandırmasında yaygın olan döndürme ve dikey çevirme gibi geometrik dönüşümler, spektrogramın yatay ekseninin zamanı, dikey ekseninin ise frekansı temsil etmesi nedeniyle sinyalin zaman-frekans yapısını fiziksel olarak anlamsız hâle getirmekte; Zhou ve diğerleri (Zhou vd. 2022) tarafından “spektrogram kısıtı” olarak adlandırılan bu durum, model başarımı üzerinde bozucu bir etki doğurmaktadır. Buna karşılık ses verilerine toplamsal Gauss gürültüsü eklenmesi, zaman-frekans yapısını koruduğu ve sinyalin fiziksel anlamını bozmadığı için, bu çalışmada segment uzunluğunun değiştirilmesiyle birlikte iki temel artırma stratejisinden biri olarak benimsenmiştir.

Deneyisel bulgular, segment uzunluğu ile doğrulama doğruluğu arasındaki ilişkinin tekdüze (monoton) olmadığını ve tepe noktası 3 saniyede bulunan bir örüntü sergilediğini ortaya koymuştur.  $N(0, 5^2)$  rejiminde ortalama doğrulama doğruluğu, segment uzunluğu 10 saniyeden 3 saniyeye indikçe  $\%85,92 \rightarrow \%86,69 \rightarrow \%88,52$  seyriyle yükselmiş, pencerenin 1 saniyeye daraltılmasıyla  $\%86,02$  düzeyine gerilemiştir.  $N(3, 5^2)$  rejiminde de aynı örüntü korunmuş; doğruluk  $\%77,96 \rightarrow \%79,67 \rightarrow \%80,51$  seyriyle yine 3

saniyede tepe yapmış ve 1 saniyede %78,62 düzeyine düşmüştür. Optimum segment uzunluğunun, birbirinden bağımsız iki gürültü rejiminde de değişmeksizin 3 saniyede kalması, bu sonucun tek bir ikili karşılaştırmaya değil tutarlı bir örüntüye dayandığını göstermektedir.

Segment uzunluğu kısaldıkça örneklem hacmi büyümüş (10 saniyede 1797, 5 saniyede 3644, 3 saniyede 6126 ve 1 saniyede 18524 spektrogram) ve katlamalar arası standart sapma her iki rejimde de tekdüze biçimde azalmıştır;  $N(0, 5^2)$  rejiminde bu azalış  $\pm\%3,14 \rightarrow \pm\%1,47 \rightarrow \pm\%1,39 \rightarrow \pm\%0,72$  biçiminde gerçekleşmiştir. Ancak en düşük standart sapmayı veren 1 saniyelik konfigürasyon en yüksek doğruluğu sağlamamış; ortalama doğruluğu, 3 saniyelik optimumun 2,50 puan altında (%88,52'ye karşı %86,02) ve neredeyse on kat daha az örnek üreten 10 saniyelik konfigürasyonunkiyile (%85,92) özdeş bir düzeyde kalmıştır. Bu durum, örneklem hacminin sağladığı istatistiksel kararlılığın tek başına yüksek başarımın güvencesi olmadığını ortaya koymaktadır. 1 saniyelik pencere, gemi kaynaklı akustik imzayı tanımlayan pervane dönüş döngüsü, makine harmonikleri ve baskın tonal bileşenler gibi periyodik yapıları tek bir spektrogramda yeterli bağlamla temsil edemediğinden, modeli yüksek bir yanlılıkla (bias) sınırlandırmaktadır. Buna karşılık 3 saniyelik pencere, söz konusu periyodik bileşenleri temsil etmeye yetecek zaman bağlamını korurken genelleme için gerekli örnek çeşitliliğini de sağlayarak iki uç arasındaki en elverişli dengeyi kurmuştur.

Gürültü rejiminin başarım üzerindeki ana etkisi incelendiğinde, ortalamada 3 dB daha yüksek gürültü ekleyerek sinyal-gürültü oranını düşüren  $N(3, 5^2)$  rejiminin, tüm segment uzunluklarında  $N(0, 5^2)$  rejiminin belirgin biçimde altında bir başarım verdiği görülmüştür. İki rejim arasındaki ortalama doğruluk farkı, 10, 5, 3 ve 1 saniyelik konfigürasyonlarda sırasıyla 7,96; 7,02; 8,01 ve 7,40 yüzde puanı olup yaklaşık 7-8 yüzde puanlık dar bir bantta seyretmiştir.

Optimum segment uzunluğunun her iki rejimde 3 saniyede sabit kalması ile iki rejim arasındaki farkın segment uzunluğundan görece bağımsız ve yaklaşık sabit kalması bir arada değerlendirildiğinde, iki tasarım değişkeninin başarım üzerindeki etkilerinin belirgin bir etkileşim içermeksizin büyük ölçüde toplamsal (additive) ve birbirinden

ayrışabilir bir karakter taşıdığı sonucuna ulaşılmıştır. Bu ayrışabilirlik, segment uzunluğu ve gürültü artırımı şiddetinin uygulamada büyük ölçüde birbirinden bağımsız olarak kalibre edilebileceğini göstermesi bakımından metodolojik bir önem taşımaktadır.

Gürültü artırımının başarıma katkısının yönü ve büyüklüğü, gürültüsüz kontrol deneyi referans alındığında çözümlenebilmiştir. 5 saniyelik segmentasyon için herhangi bir gürültü artırımı uygulanmadığında ortalama doğrulama doğruluğu %82,37 düzeyinde kalmış; aynı segment uzunluğunda  $N(0, 5^2)$  rejimi bu temel değeri 4,32 puan yükseltirken (%82,37 → %86,69),  $N(3, 5^2)$  rejimi başarıma temel değerini 2,70 puan altına düşürmüştür (%82,37 → %79,67). Bu karşılaştırma, gürültü artırımının tek yönlü ve tekdüze biçimde yararlı bir işlem olmadığını; ortalaması 0 dB olan ılımlı bir artırımın modeli gürültü örüntülerine karşı dayanıklı kılan düzenleştirici (regularizing) bir etki sağladığını, buna karşılık ortalaması 3 dB olan aşırı şiddetli bir artırımın eğitim işaretini bozarak gürültünün hiç eklenmediği duruma kıyasla dahi daha düşük bir başarıma yol açtığını ortaya koymuştur.

Sınıf bazlı başarıma örüntüleri, segment uzunluğundan ve gürültü rejiminden bağımsız, yapısal olarak kararlı bir sıralama ortaya koymuştur. Tonal hat ya da sistematik harmonik bileşen içermeyen, geniş bantlı ve stokastik yapısıyla doğal ortam gürültüsünü temsil eden Class E, tüm konfigürasyonlarda 0,93'ün üzerindeki F1-Skoru (optimum konfigürasyonda 0,9913) ile hem en yüksek ayırt edilebilirliği hem de en yüksek gürültü dayanıklılığını sergilemiştir. Kararlı çalışma rejimindeki yolcu gemilerini temsil eden ve veri kümesinde en yüksek destek değerine sahip Class D, 0,9027'ye ulaşan F1-Skoru ile ikinci sırada yer almıştır. Buna karşılık spektral içerikleri özellikle 0-4 kHz düşük-orta frekans bandında örtüşen Class A, Class B ve Class C, optimum konfigürasyonda dahi sırasıyla 0,8707, 0,8351 ve 0,8434 düzeyinde kalmış; Class B ve Class C, tüm segment uzunluklarında ve her iki rejimde en düşük başarıma ve gürültüye karşı en yüksek duyarlılığı göstermiştir. Bu sıralamanın deney koşullarından bağımsız biçimde korunması, söz konusu örüntünün deneysel bir yan üründen değil, sınıfların içsel akustik ayırt edilebilirlik derecelerinden kaynaklandığına işaret etmektedir.

Bu bulgular ışığında, ShipsEar veri seti ve EfficientNetB1 mimarisi özelinde en yüksek ortalama doğrulama doğruluğunu ( $88,52 \pm 1,39$ ) ve en yüksek makro F1-Skorunu ( $0,8887 \pm 0,0131$ ) sağlayan, aynı zamanda uzun segmentlere kıyasla belirgin biçimde düşük bir standart sapma sunan 3 saniyelik segmentasyon ile  $N(0, 5^2)$  gürültü rejiminin birlikte kullanımı, bu çalışma kapsamındaki optimal konfigürasyon olarak belirlenmiştir.

Elde edilen bulgular bütün olarak değerlendirildiğinde, çalışma su altı akustik hedef tanımda derin öğrenme modeli tasarımına ilişkin dört temel kavramsal çıkarımı desteklemektedir.

Birincisi, segment uzunluğu yalnızca örneklem hacmini belirleyen bir veri artırma parametresi değil, modele aktarılan akustik bağlamın fiziksel anlamlılığını gözetten bir optimizasyon parametresidir. Mimarinin sabit tutulmasına karşın girdi spektrogramının temsil ettiği akustik zaman penceresinin değişmesi başarımda belirgin farklar oluşturmuş; en büyük örneklem hacmini üreten 1 saniyelik konfigürasyonun yaklaşık on katlık örnek artışına rağmen optimumun altında kalması, “ne kadar çok örnek, o kadar iyi” sezgisinin, sinyallerin ayırt ediciliğinin fiziksel zaman ölçeklerine bağlı olduğu su altı akustiği gibi alanlarda geçerliliğini yitirdiğini göstermiştir. Bu çıkarım, alanın daha derin veya daha geniş mimari arayışlarından önce girdi katmanının optimizasyonuna öncelik vermesi gerektiğini önermektedir.

İkincisi, veri artırma yalnızca örneklem hacmini artıran tek yönlü bir işlem değil, dikkatle kalibre edilmesi gereken çift yönlü bir araçtır. Ilımlı bir gürültü artırımı genellemeyi iyileştiren bir düzenleştirci işlevi görürken, aşırı şiddetli bir artırım etkin sinyal-gürültü oranını düşürerek başarıyı gürültüsüz temelin dahi altına çekebilmektedir. Bu nedenle veri artırma tasarımı, artırımın yalnızca türünü değil şiddetini de sinyalin fiziksel doğasını ve etkin sinyal-gürültü oranını koruyacak biçimde belirlemelidir.

Üçüncüsü, faktöriyel tasarım, iki tasarım değişkeninin başarımlar üzerindeki etkilerinin büyük ölçüde toplamsal ve ayrışabilir olduğunu ortaya koyarak tasarım uzayının yapısına ilişkin bir çıkarım sağlamıştır. Optimum segment uzunluğunun gürültü rejiminden, gürültü rejiminin etkisinin ise segment uzunluğundan görece bağımsız kalması, bu iki

parametrenin birbirinden ayrı biçimde iyileştirilebileceğini göstermekte; aynı zamanda sabit bir mimari altında dahi girdi temsiline ilişkin kararların başarımın birincil belirleyicisi olduğunu vurgulamaktadır.

Dördüncüsü, sınıf bazlı başarım sıralamasının deney koşullarından bağımsız biçimde kararlı kalması, model değerlendirmesinin yalnızca genel doğruluk üzerinden yapılmasının yanıltıcı olabileceğini açıkça ortaya koymaktadır. Optimum konfigürasyonda Class E için 0,9913 düzeyindeki F1-Skoruna karşılık Class B için 0,8351 düzeyinde kalan başarım, sınıflar arasında niteliksel bir güvenilirlik farkına işaret etmektedir. Tek başına ortalama doğruluğu raporlamak bu farklılıkları gizlediğinden, özellikle hata maliyetinin yüksek olduğu savunma uygulamalarında kritik operasyonel sonuçlar doğurabilir. Bu nedenle hata matrisinin ve her sınıf için hesaplanan kesinlik, duyarlılık ile F1-Skoru metriklerinin birlikte raporlanması, model değerlendirmesinin asgari bir gerekliliği olarak benimsenmelidir.

## KAYNAKLAR

- Anonymous. (2025a). *Dosits-Small Vessels*. <https://dosits.org/animals/effects-of-sound/anthropogenic-sources/small-vessels/>
- Anonymous. (2025b). *Multibeam Echosounder*. <https://dosits.org/galleries/technology-gallery/observing-the-sea-floor/multibeam-echosounder/>
- Aslam, M. A., Zhang, L., Liu, X., Irfan, M., Xu, Y., Li, N., Zhang, P., Jiangbin, Z., & Yaan, L. (2024). Underwater sound classification using learning based methods: A review. *Expert Systems with Applications*, 255. <https://doi.org/10.1016/j.eswa.2024.124498>
- Doan, V. S., Huynh-The, T., & Kim, D. S. (2022). Underwater Acoustic Target Classification Based on Dense Convolutional Neural Network. *IEEE Geoscience and Remote Sensing Letters*, 19. <https://doi.org/10.1109/LGRS.2020.3029584>
- Dörfler, M., Bammer, R., & Grill, T. (2017). *Inside the Spectrogram: Convolutional Neural Networks in Audio Processing*. 152-155.
- Du, X., & Hong, F. (2024). *QiandaoEar22: A high quality noise dataset for identifying specific ship from multiple underwater acoustic targets using ship-radiated noise*. <http://arxiv.org/abs/2406.04354>
- Ghavidel, M., Azhdari, S. M. H., Khishe, M., & Kazemirad, M. (2022). Sonar data classification by using few-shot learning and concept extraction. *Applied Acoustics*, 195. <https://doi.org/10.1016/j.apacoust.2022.108856>
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*. <http://arxiv.org/abs/1512.03385>
- Hong, F., Liu, C., Guo, L., Chen, F., & Feng, H. (2021). Underwater acoustic target recognition with a residual network and the optimized feature extraction method. *Applied Sciences (Switzerland)*, 11(4), 1-12. <https://doi.org/10.3390/app11041442>
- Hossan, Md. A., Memon, S., & Gregory, M. A. (2010). A Novel Approach for MFCC Feature Extraction. *2010 4th International Conference on Signal Processing and Communication Systems*. IEEE.
- Kaharl, V. (1999). *Sounding Out the Ocean's Secrets*. [https://misclab.umeoce.maine.edu/boss/classes/SMS\\_598\\_2005/PDFs/Sounding%20Out%20the%20Ocean's%20Secrets.pdf](https://misclab.umeoce.maine.edu/boss/classes/SMS_598_2005/PDFs/Sounding%20Out%20the%20Ocean's%20Secrets.pdf)

- Lee, S., Seo, I., Seok, J., Kim, Y., & Han, D. S. (2020). Active sonar target classification with power- normalized cepstral coefficients and convolutional neural network. *Applied Sciences (Switzerland)*, 10(23), 1-15. <https://doi.org/10.3390/app10238450>
- Liu, F., Shen, T., Luo, Z., Zhao, D., & Guo, S. (2021). Underwater target recognition using convolutional recurrent neural networks with 3-D Mel-spectrogram and data augmentation. *Applied Acoustics*, 178. <https://doi.org/10.1016/j.apacoust.2021.107989>
- National Oceanic and Atmospheric Administration. (2024a, Haziran 16). *What is SOFAR?* <https://oceanservice.noaa.gov/facts/sofar.html>
- National Oceanic and Atmospheric Administration. (2024b, Haziran 16). *What is sonar?* <https://oceanservice.noaa.gov/facts/sonar.html>
- National Oceanic and Atmospheric Administration. (2024c, Ağustos 29). *Synthetic Aperture Sonar (SAS)*. 2024. <https://oceanexplorer.noaa.gov/technology/sonar-sas/>
- Neupane, D., & Seok, J. (2020). A review on deep learning-based approaches for automatic sonar target recognition. İçinde *Electronics (Switzerland)* (C. 9, Sayı 11, ss. 1-30). MDPI AG. <https://doi.org/10.3390/electronics9111972>
- Santos-Domínguez, D., Torres-Guijarro, S., Cardenal-López, A., & Pena-Gimenez, A. (2016). ShipsEar: An underwater vessel noise database. *Applied Acoustics*, 113, 64-69. <https://doi.org/10.1016/j.apacoust.2016.06.008>
- Shen, S., Yang, H., Li, J., Xu, G., & Sheng, M. (2018). Auditory inspired convolutional neural networks for ship type classification with raw hydrophone data. *Entropy*, 20(12). <https://doi.org/10.3390/e20120990>
- Shipton, M., Obradović, J., Ferreira, F., Mišković, N., Bulat, T., Cukrov, N., & Diamant, R. (2025). A Database of Underwater Radiated Noise from Small Vessels in the Coastal Area. *Scientific Data* , 12(1). <https://doi.org/10.1038/s41597-025-04584-x>
- Tan, M., & Le, Q. V. (2019). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. <http://arxiv.org/abs/1905.11946>
- Tian, S., Chen, D., Wang, H., & Liu, J. (2021). Deep convolution stack for waveform in underwater acoustic target recognition. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-88799-z>
- Waite, A. (2002). *Sonar for Practising Engineers Third Edition*.
- Xu, J., Xie, Y., & Wang, W. (2024). *Underwater Acoustic Target Recognition based on Smoothness-inducing Regularization and Spectrogram-based Data Augmentation*. <https://doi.org/10.1016/j.oceaneng.2023.114926>

- Yang, S., Jin, A., Zeng, X., Wang, H., Hong, X., & Lei, M. (2024). Underwater acoustic target recognition based on sub-band concatenated Mel spectrogram and multidomain attention mechanism. *Engineering Applications of Artificial Intelligence*, 133. <https://doi.org/10.1016/j.engappai.2024.107983>
- Yao, Q., Wang, Y., & Yang, Y. (2023). Underwater Acoustic Target Recognition Based on Data Augmentation and Residual CNN. *Electronics (Switzerland)*, 12(5). <https://doi.org/10.3390/electronics12051206>
- Zeng, D., Yan, S., Yang, J., & Pan, X. (2025). An efficient deep learning approach with frequency and channel optimization for underwater acoustic target recognition. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-12452-2>
- Zhou, G., Chen, Y., & Chien, C. (2022). On the analysis of data augmentation methods for spectral imaged based heart sound classification using convolutional neural networks. *BMC medical informatics and decision making*, 22(1), 226. <https://doi.org/10.1186/s12911-022-01942-2>
- Zhu, C., Gaggero, T., Makris, N. C., & Ratilal, P. (2022). Underwater Sound Characteristics of a Ship with Controllable Pitch Propeller. *Journal of Marine Science and Engineering*, 10(3). <https://doi.org/10.3390/jmse10030328>