



**ESKİŞEHİR OSMANGAZI ÜNİVERSİTESİ**  
**MÜHENDİSLİK-MİMARLIK FAKÜLTESİ**  
**BİLGİSAYAR MÜHENDİSLİĞİ**

**NEURAL NETWORK ANOMALİ TESPİTİ PROJESİ**

**Ders Yürütücüsü:**  
**Doç. Dr. Eyüp ÇINAR**

**Hazırlayan:**  
**Doğukan KIYIKLIK – 152120211104**

**Mayıs 2025**

# İçindekiler

1. Giriş .....	3
2. Kullanılan Yöntemler ve Modeller .....	3
2.1. Ortak Modelleme Yapısı ve Veri Hazırlığı .....	3
2.2. Değerlendirilen Modeller .....	4
3. Uygulama Adımları .....	4
3.1. Ortam Kurulumu ve Arayüzün Başlatılması: .....	4
3.2. Model Eğitimi ve Test İşlemleri: .....	4
3.3. Sonuçların İncelenmesi ve Değerlendirilmesi: .....	5
4. Sonuçlar ve Performans Değerlendirmesi .....	5
4.1 Sayısal Performans Metrikleri ve Karşılaştırmaları .....	6
4.2 Loss Grafikleri .....	7
4.3. Sonuçların Genel Analizi ve Karşılaşılan Zorluklar .....	8
4.3.1. Sınıflandırma Başarısı ve Segmentasyon Zorluğu Dengesi .....	8
4.3.2. Overfitting Durumunun İncelenmesi .....	8
4.3.3. Feature Extractor Tabanlı Modellerin (PaDiM ve PatchCore) Performansı .....	8
4.3.4. Genel Çıkarım .....	8

# 1. Giriş

Ahşap yüzeylerdeki üretim hatalarının tespiti, kalite kontrol süreçlerinin otomasyonu açısından kritik öneme sahiptir. Bu projede, ahşap yüzey görüntüleri üzerinde anomali tespiti ve segmentasyonu yapmak amacıyla denetimsiz öğrenme yaklaşımları kullanılmıştır. Kullanılan yöntemler, doğrudan etiketli anomali bilgisine ihtiyaç duymadan, yalnızca normal görüntülerden öğrenerek anomalili bölgeleri belirlemeyi hedeflemiştir.

Proje kapsamında EfficientAD, PaDiM, PatchCore, STFCM, DRAEM ve Reverse Distillation olmak üzere toplam altı farklı derin öğrenme tabanlı model incelenmiş ve karşılaştırılmıştır. Bu modeller, segmentasyon maskeleri ve anomali skorları üretme yeteneklerine göre değerlendirilmiştir. Modellerin performansı hem sınıflandırma doğruluğu (F1 Score) hem de segmentasyon kalitesi (IoU – Intersection over Union) açısından ölçülmüştür.

Çalışmada kullanılan ahşap veri seti, yalnızca "wood" sınıfını içermektedir. Veri seti; eğitim (train/good), test (test/good ve test/defect) ve gerçek segmentasyon maskelerini (ground\_truth/defect) içeren yapısıyla, modellerin anomali tespit kabiliyetlerini değerlendirmek için uygundur. Tüm görüntüler ön işlem adımlarından geçirilmiş, 256x256 piksele yeniden boyutlandırılmıştır.

Projenin bir diğer hedefi, bu modellerin test sonuçlarının kullanıcı dostu bir arayüz üzerinden görselleştirilmesidir. Bu amaçla Gradio tabanlı bir kullanıcı arayüzü geliştirilmiş; arayüzde model eğitimi, testi, maskelerin görselleştirilmesi, performans karşılaştırmaları ve loss eğrilerinin gösterimi gibi işlevler birleştirilmiştir. Tüm proje, tekrarlanabilirlik ilkesine uygun olarak açık kaynaklı bir GitHub reposunda ve Google Colab üzerinde çalıştırılabilir formatta sunulmuştur.

## 2. Kullanılan Yöntemler ve Modeller

Bu çalışmada, ahşap yüzeylerdeki üretim hatalarını tespit etmek amacıyla ahşap veri seti üzerinde denetimsiz anomali tespiti yapan altı farklı derin öğrenme modeli değerlendirilmiştir. Tüm modeller, anomali tespiti için standart bir çerçeve sunan Anomalib kütüphanesi temel alınarak entegre edilmiştir. Modellerin hem anomali skoru hesaplama yetenekleri hem de anomali bölgelerini işaretleyen segmentasyon maskeleri üretme başarıları test edilmiştir.

### 2.1. Ortak Modelleme Yapısı ve Veri Hazırlığı

Denetimsiz öğrenme prensibine uygun olarak tüm modeller yalnızca kusursuz ahşap görüntüleri kullanılarak eğitilmiştir. Bu sayede, modellerin daha önce hiç karşılaşmadıkları kusurlu durumları öğrenilmiş normal örüntülerden sapma olarak tespit etmesi hedeflenmiştir.

Veri ön işleme aşamasında, preprocessed.py scripti ile otomatikleştirilmiş adımlar izlenmiştir:

1. Görüntüler öncelikle en büyük kontur baz alınarak kırpılmış,
2. Ardından tüm görüntüler 256×256 piksel boyutuna yeniden boyutlandırılmış,
3. Özellikle eğitimde kullanılan train/good klasöründeki normal görüntülere, modelin genelleme yeteneğini artırmak amacıyla yatay ve dikey çevirme augmentasyonları uygulanmıştır.

Bu standartlaştırılmış veri hazırlık süreci, tüm modellerin adil bir şekilde karşılaştırılabilmesi için kritik öneme sahiptir.

## 2.2. Değerlendirilen Modeller

Proje kapsamında incelenen ve karşılaştırılan modeller şunlardır:

- EfficientAD
- PaDiM
- PatchCore
- STFPM
- DRAEM
- Reverse Distillation

Modellere ait F1 ve IoU skorları, eğitim ve test süreçlerinin çıktısı olan log dosyalarından `gradio_arayuz.py` içerisindeki `run_and_parse()` fonksiyonu aracılığıyla otomatik olarak çekilmiş ve kullanıcı dostu bir Gradio arayüzünde karşılaştırmalı bir tablo olarak sunulmuştur. Bu otomasyon ve arayüz entegrasyonu, projenin deneysel sonuçlarının kolayca incelenmesine ve tekrar üretilebilirliğine katkı sağlamıştır.

## 3. Uygulama Adımları

Bu bölümde, geliştirilen ahşap yüzey anomali tespit sisteminin nasıl çalıştırılacağı ve sonuçların nasıl inceleneceği adım adım açıklanmaktadır. Projenin tamamı, erişilebilirliği ve tekrarlanabilirliği en üst düzeye çıkarmak amacıyla, tüm kurulum ve çalışma adımlarını içeren bir Google Colab not defteri (Kullanım\_Kılavuzu.ipynb) ile desteklenmiş ve GitHub üzerinden açık kaynak olarak paylaşılmıştır.

### 3.1. Ortam Kurulumu ve Arayüzün Başlatılması:

- **Adım 1:** Projenin GitHub repositorisine erişiniz.
- **Adım 2:** Repositorinin ana dizininde bulunan `Kullanım_Kılavuzu.ipynb` isimli Google Colab dosyasını açınız.
- **Adım 3:** Colab not defterindeki hücreleri, dosyanın başından sonuna doğru sırasıyla çalıştırınız. Bu işlem, gerekli Python kütüphanelerinin kurulmasını, projenin bağımlılıklarının yüklenmesini ve Gradio tabanlı interaktif kullanıcı arayüzünün başlatılmasını sağlayacaktır.

### 3.2. Model Eğitimi ve Test İşlemleri:

- **Adım 1:** Colab hücrelerinin tamamı başarıyla çalıştırıldığında, hücre çıktılarının sonunda Gradio arayüzüne ait bir public URL adresi belirecektir. Bu linke tıklayarak interaktif panele tarayıcınız üzerinden ulaşınız.
- **Adım 2:** Arayüz açıldığında, "Eğitim / Test Paneli" sekmesine gidiniz. Bu sekmede, denemek istediğiniz anomali tespit modelini (EfficientAD, PaDiM, PatchCore, STFPM, DRAEM, veya Reverse Distillation) ilgili açılır menüden seçiniz.

- **Adım 3:** Model seçiminin ardından, "İşlem Türü" bölümünden "Train" veya "Test" seçeneğini işaretleyiniz.
  - "Train" seçeneği, seçilen modeli preprocessed\_wood\_dataset/wood/train/good/ altındaki normal verilerle eğitecek ve eğitilmiş model ağırlıklarını checkpoints/ klasörü altına kaydedecektir. Eğitim süreciyle ilgili loglar logs/ klasöründe tutulacaktır.
  - "Test" seçeneği, daha önce eğitilmiş ve checkpoints/ altında ağırlıkları bulunan bir modeli kullanarak preprocessed\_wood\_dataset/wood/test/ altındaki görüntüler üzerinde anomali tespiti yapacaktır.
- **Adım 4:** "Başlat" butonuna tıklayarak seçili işlemi başlatınız.

### 3.3. Sonuçların İncelenmesi ve Değerlendirilmesi:

- Eğitim veya test işlemi tamamlandıktan sonra, elde edilen sonuçlar Gradio arayüzündeki ilgili sekmelerde otomatik olarak güncellenerek kullanıcıya sunulur:
  - **"Test Sonuçları" Sekmesi:** Seçilen modelin test veri seti üzerindeki performansını gösteren sayısal metrikler (F1 Skoru, IoU, AUROC vb.) tablo halinde sunulur. Bu metrikler, logs/ klasöründeki test loglarından okunur.
  - **"Segmentasyon Maskeleri" Sekmesi:** Test edilen kusurlu ve normal görüntülerden rastgele örnekler, bu görüntülere ait orijinal haller, gerçek anomali maskeleri ve modelin ürettiği anomali segmentasyon maskeleri yan yana görselleştirilir.
  - **"Model Karşılaştırması" Sekmesi:** Test edilen tüm modellerin temel performans metriklerini içeren özet bir karşılaştırma tablosu sunulur. Modellerin birbirlerine göre başarımları kolayca değerlendirilebilir.
  - **"Eğitim Kaybı (Loss) Grafiği" Sekmesi:** "Train" işlemi yapılan bir modelin eğitim süreci boyunca her bir epoktaki loss değerlerinin değişimini gösteren bir grafik çizilir. Bu grafik, logs/ klasöründeki eğitim loglarından yararlanılarak oluşturulur.

## 4. Sonuçlar ve Performans Değerlendirmesi

Bu bölümde, altı farklı denetimsiz anomali tespit modelinin (EfficientAD, PaDiM, PatchCore, STFCM, DRAEM ve Reverse Distillation) ahşap yüzey veri seti üzerindeki performansları sunulmakta ve karşılaştırmalı olarak değerlendirilmektedir. Modellerin başarısı, iki temel metrik üzerinden ölçülmüştür:

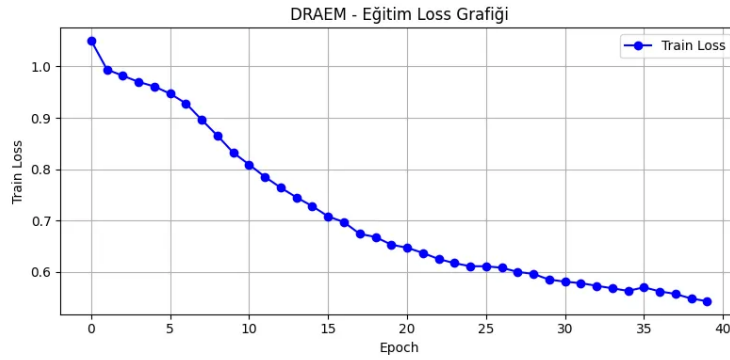
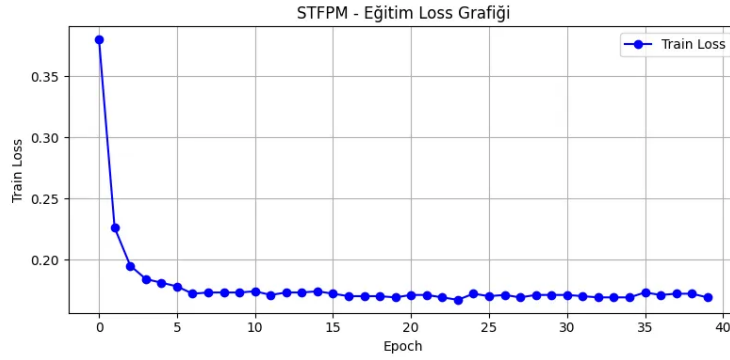
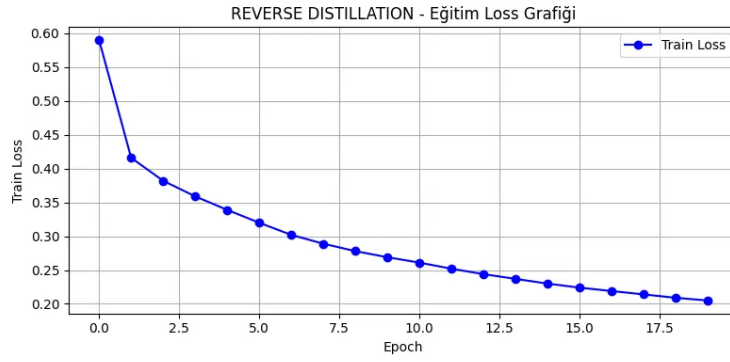
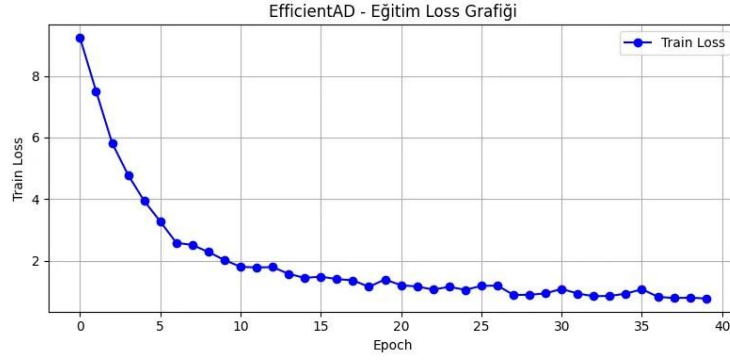
- **F1 Skoru:** Görüntü düzeyinde anomali sınıflandırma performansını değerlendirmek için kullanılmıştır.
- **IoU (Intersection over Union):** Piksel düzeyinde anomali segmentasyon kalitesini, yani model tarafından üretilen anomali maskelerinin gerçek anomali bölgeleriyle ne kadar örtüştüğünü ölçmek için kullanılmıştır.

Ek olarak, AUROC gibi diğer önemli metrikler ve modellerin eğitim süreçleri de analize dahil edilmiştir. Tüm sayısal sonuçlar, geliştirilen Gradio arayüzü aracılığıyla elde edilmiş ve aşağıda detaylandırılmıştır.

#### 4.1 Sayısal Performans Metrikleri ve Karşılaştırmaları

Model	Image AUROC	Image F1 Score	Pixel AUROC	Pixel F1Score	IoU
EfficientAD	0.9095	0.8358	0.8143	0.3844	0.302
PaDiM	0.7912	0.7654	0.9400	0.3387	0.170
PatchCore	0.8206	0.7187	0.8987	0.1985	0.159
STFPM	0.8642	0.7894	0.9084	0.3852	0.149
DRAEM	0.9043	0.8333	0.6299	0.0776	0.048
REVERSE DISTILLATION	0.9023	0.8607	0.9385	0.3432	0.167

## 4.2 Loss Grafikleri



### 4.3. Sonuçların Genel Analizi ve Karşılaşılan Zorluklar

#### 4.3.1. Sınıflandırma Başarısı ve Segmentasyon Zorluğu Dengesi

Modellerin birçoğu bir ahşap görüntüsünün genel olarak kusurlu olup olmadığını belirlemede %83-%86 aralığında F1 skorları ile başarılı bir sınıflandırma performansı sergilemiştir. Bu, modellerin normal ve anormal görüntüler arasında genel bir ayırım yapabildiğini göstermektedir. Ancak, bu sınıflandırma başarısı aynı ölçüde segmentasyon performansına yansımamıştır. En yüksek IoU skorunu %30.2 ile EfficientAD elde ederken, diğer modellerin IoU değerleri bu skora oldukça altındadır. Bu durum, modellerin bir görüntüde anomali varlığını tespit etmede görece başarılı olsalar da bu anomalinin tam olarak nerede olduğunu ve sınırlarını piksel hassasiyetinde belirlemede zorlandıklarını göstermektedir. Farklı hiperparametre denemelerine ve optimizasyon çabalarına rağmen IoU skorlarında beklenen seviyede bir iyileşme sağlanamamıştır.

#### 4.3.2. Overfitting Durumunun İncelenmesi

Düşük segmentasyon performansının olası nedenlerinden biri olarak modellerin aşırı öğrenmesi (overfitting) düşünülebilir. Ancak Bölüm 4.2’de sunulan eğitim kaybı (loss) grafikleri incelendiğinde tüm modeller için eğitim kaybının epok boyunca genel bir düşüş eğilimi gösterdiği ve belirgin bir overfitting durumuna işaret eden bir örüntü gözlemlenmemiştir. Bu, modellerin eğitim verisindeki normal örüntüleri öğrendiğini, ancak bu öğrenmenin karmaşık ahşap dokusu üzerindeki ince ve değişken anomali sınırlarını hassas bir şekilde ayırt etmek için yeterli olmayabileceğini düşündürmektedir.

#### 4.3.3. Feature Extractor Tabanlı Modellerin (PaDiM ve PatchCore) Performansı

PaDiM ve PatchCore modelleri, önceden eğitilmiş derin özellik çıkarıcı ağlara dayanır ve genellikle tek bir eğitim epokunda özellik dağılımlarını ya da bellek bankalarını oluştururlar. Bu çalışmada, bu iki modelin hem sınıflandırma hem de özellikle segmentasyon metriklerinde diğer modellere kıyasla daha düşük performans sergilediği görülmüştür. Bunun olası nedenleri arasında, kullanılan özellik çıkarıcının ahşap dokusundaki ince anomali örüntüleri için optimal özellikler üretememesi veya bu modellerin doğası gereği daha karmaşık yeniden yapılandırma veya karşılaştırma mekanizmalarına sahip modellere göre segmentasyon detaylarında daha az hassas olmaları sayılabilir. Tek epokluk eğitimleri, hızlı bir model oluşturma avantajı sunsa da karmaşık dokular ve değişken anomali tipleri karşısında segmentasyon doğruluğu açısından bir sınırlama getirmiş olabilir.

#### 4.3.4. Genel Çıkarım

İncelenen modeller, bir anomali varlığını tespit etme konusunda umut verici sonuçlar verse de, bu anomalilerin kesin sınırlarını belirlemede (IoU ile ölçülen) iyileştirme alanları bulunmaktadır. EfficientAD, hem sınıflandırma hem de segmentasyonda en dengeli performansı sunarken, Reverse Distillation sınıflandırmada öne çıkmıştır. Düşük IoU skorları, sadece model seçiminin değil, aynı zamanda veri ön işleme, özellik temsili, eşikleme stratejileri ve model hiperparametrelerinin de segmentasyon kalitesi üzerinde kritik bir etkiye sahip olduğunu göstermektedir.



