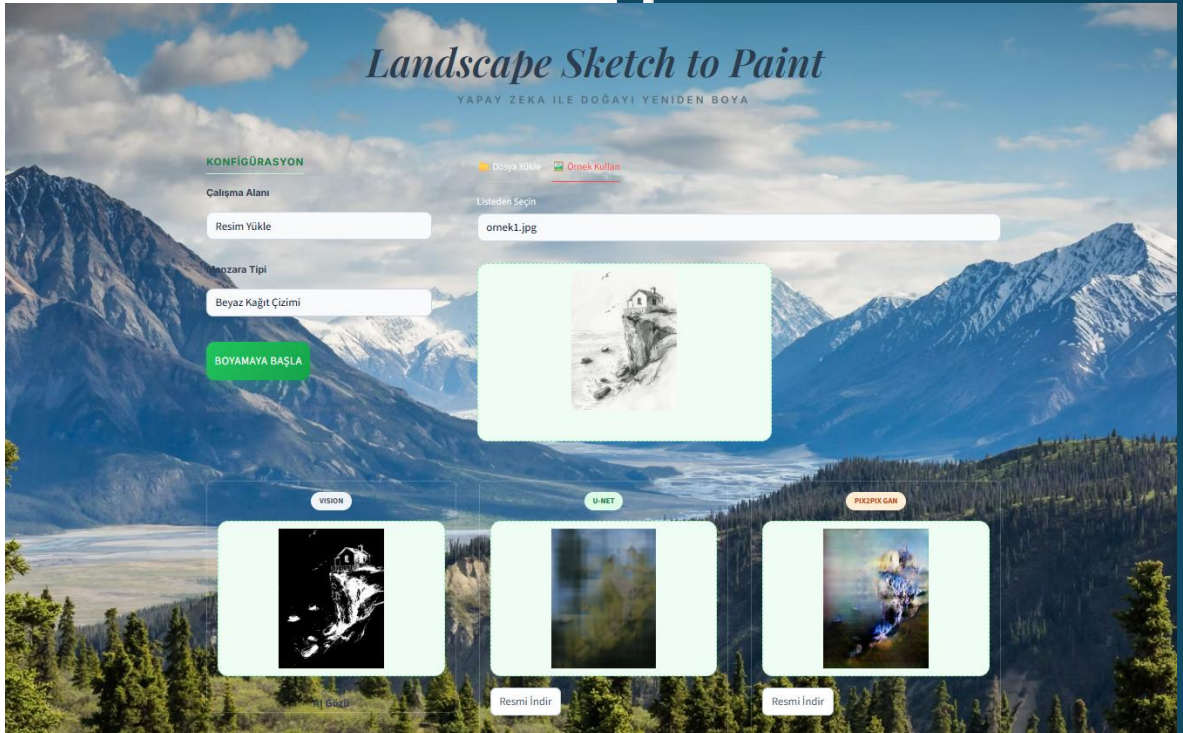


BIL485-DeepLearning Proje Ödevi



Hazırlayan

Fatımanur Kantar, 23120205089

Öğretim Üyesi

Dr. Öğr. Üyesi İshak DÖLEK

30.12.2025

İçindekiler

Özet (Abstract)	2
1. Giriş (Introduction)	3
2. Materyal ve Yöntem	4
2.1. Veri Seti ve Dinamik Ön İşleme	4
2.2. Ağ Mimarileri	4
2.2.1. Baseline Model (Standart U-Net)	4
2.2.2. Önerilen Model: Pix2Pix (cGAN)	5
2.3. Kayıp Fonksiyonları (Loss Functions)	6
2.3.1. Adversarial Loss (Çekişmeli Kayıp)	6
2.3.2. L1 Loss (Piksel Bazlı Kayıp)	7
2.3.3. Toplam Amaç Fonksiyonu	7
2.4. Deneysel Kurgu	7
2.5. Eğitim Donanımı ve Hiperparametreler	8
3. Deneysel Bulgular	9
3.1. Eğitim Kararlılık Analizi	9
3.2. Görsel Karşılaştırma ve Nitel Analiz	10
3.2.1. Baseline U-Net Sonuçları	10
3.2.2. Pix2Pix GAN Sonuçları	11
3.3. Sayısal Performans Değerlendirmesi	12
4. Sonuç	13
Kaynakça	14

Özet (Abstract)

Bu çalışmada, bilgisayarlı görü alanında önemli bir problem olan görüntüden görüntüye çeviri (Image-to-Image Translation) konusu ele alınmıştır. Çalışmanın temel amacı, anlamsal etiket haritalarını veya kenar taslaklarını (sketch), fotorealistik manzara görüntülerine dönüştürmektir. Bu kapsamda, piksel tabanlı kayıp fonksiyonlarını kullanan geleneksel Konvolüsyonel Sinir Ağları (CNN) ile hem piksel hem de algısal (perceptual) kaybı optimize eden Koşullu Üretken Çekişmeli Ağlar (cGAN / Pix2Pix) karşılaştırılmıştır. Eğitim süreçleri NVIDIA Tesla T4 GPU üzerinde gerçekleştirilmiş olup, dinamik veri üretimi (runtime sketch generation) yöntemiyle modelin dayanıklılığı artırılmıştır. Deneysel sonuçlar, L1 kaybı kullanan U-Net modelinin düşük frekanslı detayları öğrendiğini ancak bulanık görüntüler ürettiğini; PatchGAN diskriminatörüne sahip Pix2Pix modelinin ise yüksek frekanslı detayları (doku, keskinlik) öğrenerek insan algısına daha uygun sonuçlar ürettiğini göstermiştir.

Anahtar Kelimeler: *Derin Öğrenme, GAN, Pix2Pix, U-Net, Görüntü Sentezi, Bilgisayarlı Görü.*

1. Giriş (Introduction)

Görüntü sentezi ve görüntüden görüntüye çeviri (Image-to-Image Translation), bir giriş görüntüsünün anlamsal içeriğini koruyarak farklı bir alana (domain) transfer edilmesi işlemidir. Bu problem, siyah-beyaz çizimlerin renklendirilmesinden, uydu görüntülerinin haritaya dönüştürülmesine kadar geniş bir uygulama alanına sahiptir. Ancak, özellikle serbest el çizimi (freehand sketch) gibi düşük bilgi içeren girdilerden fotorealistik görüntü üretmek, "ill-posed" (iyi tanımlanmamış) bir problemdir. Çünkü tek bir çizgi, gerçek dünyada birçok farklı nesneye veya dokuya karşılık gelebilir.

Bu alandaki çalışmalarda uzamsal bilginin korunması kritik öneme sahiptir. Bu bağlamda, Ronneberger ve arkadaşları [2] tarafından geliştirilen **U-Net** mimarisi, kodlayıcı ve kod çözücü katmanları arasında kurduğu "atlamalı bağlantılar" (skip connections) sayesinde, giriş görüntüsündeki yapısal sınırların kaybolmadan çıkışa aktarılmasını sağlamış ve bir standart haline gelmiştir. Ancak, geleneksel CNN yapıları (U-Net dahil), eğitim sırasında piksel farklarını (L1 veya L2 normu) minimize etmeye odaklandığında, belirsizlik durumlarında olası renklerin ortalamasını alma eğilimi gösterir. Bu durum, üretilen görsellerin dokusal açıdan zayıf ve "bulanık" (blurry) olmasına neden olmaktadır.

Söz konusu bulanıklık problemini aşmak ve gerçekçi doku sentezi sağlamak amacıyla, literatürdeki odak noktası Üretken Çekişmeli Ağlara (GAN) kaymıştır. Goodfellow ve arkadaşları [1] tarafından önerilen GAN yapısı, veri dağılımını öğrenen bir üretici ve bu üretimin gerçekliğini sorgulayan bir ayırt edici ağın rekabetçi oyun teorisi çerçevesinde eğitilmesi prensibine dayanır. Isola ve arkadaşları [3] ise bu yapıyı görüntüden görüntüye çeviri görevleri için özelleştirerek **Pix2Pix** mimarisini önermiştir. Pix2Pix, L1 kaybını (düşük frekanslı doğruluğu sağlar) ve çekişmeli kaybı (yüksek frekanslı keskinliği sağlar) hibrit bir şekilde kullanarak literatürde devrim yaratmıştır.

Bu proje kapsamında, Isola ve arkadaşlarının önerdiği yaklaşım temel alınmış ve dinamik veri ön işleme (dynamic preprocessing) yöntemleri kullanılarak sınırlı veri setlerinde model genelleştirmesinin artırılması hedeflenmiştir. Çalışmada iki aşamalı bir deney kurgusu izlenmiştir:

1. **Baseline Model:** Sadece L1 kaybı kullanan U-Net tabanlı bir model.
2. **Önerilen Model:** Girdi görüntüsünü şart (condition) olarak alan ve bir "PatchGAN" diskriminatörü ile eğitilen Pix2Pix mimarisi.

Projenin temel katkısı, U-Net ile GAN mimarileri arasındaki performans farkını hem görsel hem de nicel (SSIM, PSNR) metriklerle ortaya koymaktır.

2. Materyal ve Yöntem

Bu bölümde, önerilen yöntemin veri hazırlama süreçleri, kullanılan ağ mimarileri, kayıp fonksiyonlarının matematiksel formülasyonu ve eğitim detayları sunulmuştur.

2.1. Veri Seti ve Dinamik Ön İşleme

Çalışmada Kaggle platformunda erişime açık olan "Landscape Pictures" veri seti kullanılmıştır. Ancak, modelin genelleştirme yeteneğini artırmak ve farklı çizim stillerine karşı dayanıklılık (robustness) sağlamak amacıyla statik bir eğitim seti yerine **dinamik veri üretimi (runtime data generation)** yöntemi benimsenmiştir.

Eğitim sürecinin her iterasyonunda, orijinal RGB görüntüler gri seviyeye dönüştürülmüş ve **Canny Kenar Tespiti (Canny Edge Detection)** algoritması uygulanarak taslak (sketch) görüntüleri çalışma zamanında üretilmiştir. Canny algoritması için alt ve üst eşik değerleri (thresholds) dinamik olarak belirlenmiş, böylece modelin farklı çizgi kalınlıklarına ve detay seviyelerine adapte olması sağlanmıştır.

Elde edilen taslak ve gerçek görüntü çiftleri, modelin aktivasyon fonksiyonuna uygun olacak şekilde $[0, 1]$ aralığına normalize edilmiş ve 256x256 piksel boyutuna yeniden ölçeklendirilmiştir.

2.2. Ağ Mimarileri

Bu çalışmada, taslak çizimlerin renklendirilmesi problemi için iki farklı derin öğrenme mimarisi kurulmuş ve analiz edilmiştir. İlk olarak referans noktası oluşturması amacıyla tek başına çalışan bir **Baseline U-Net**, ardından önerilen yöntem olan **Koşullu Üretken Çekişmeli Ağ (Pix2Pix cGAN)** mimarisi geliştirilmiştir.

2.2.1. Baseline Model (Standart U-Net)

Referans model olarak, görüntüdeki uzamsal yapıyı koruma başarısı nedeniyle **U-Net** mimarisi tercih edilmiştir. U-Net, bir kodlayıcı (encoder) ve simetrik bir kod çözücünden (decoder) oluşur. Kodlayıcı katmanlarında görüntünün boyutları küçültülerek öznetelikler çıkarılırken, kod çözücü katmanlarında görüntü tekrar orijinal boyutuna getirilir.

Standart yapılardan farklı olarak, U-Net'te **atlamalı bağlantılar (skip connections)** bulunur. Bu bağlantılar, kodlayıcıdaki i . katmanın çıktısını, kod çözücüdeki $(n-i)$. katman ile birleştirerek (concatenation), darboğaz (bottleneck) katmanında kaybolan düşük seviyeli detayların geri kazanılmasını sağlar.

Mimari Uygulama Detayları:

Blok Tipi	Katman İşlemleri	Çıkış Boyutu	Filtre Sayısı
Giriş	Input Layer	256×256	1 (Gray)
Encoder	4 Blok (Conv2D + BatchNorm + ReLU)	256→16	64→512
Bottleneck	MaxPool + Conv2D + BatchNorm	16×16	1024
Decoder	4 Blok (ConvTranspose + Concat)	32→256	512→64
Çıkış	Conv2D (Sigmoid/Tanh)	256×256	3 (RGB)

Tablo 1: Baseline U-Net Modelinin Mimari Özeti

- **Toplam Parametre:** 31,054,275
- **Eğitilebilir Parametre:** 31,042,499

2.2.2. Önerilen Model: Pix2Pix (cGAN)

Önerilen ana yöntem, bir **Üretici (Generator)** ve bir **Ayırt Edici (Discriminator)** ağından oluşan Pix2Pix mimarisidir.

A) Üretici (Generator) Mimarisi: Pix2Pix içerisindeki üretici ağ da temelde U-Net yapısındadır. Ancak Baseline modelden farklı olarak, daha karmaşık dokuları öğrenebilmesi için **daha derin ve yüksek kapasiteli** tasarlanmıştır. Modelin toplam parametre sayısı yaklaşık 54 milyondur (Tablo 2).

Blok Yapısı	İşlem	Çıkış Boyutu	Açıklama
Encoder	8 Blok (Conv2D + LeakyReLU)	256→1	Görüntü 1x1 boyutuna kadar sıkıştırılır.
Bottleneck	En Derin Katman	1×1	512 Filtre
Decoder	8 Blok (ConvTranspose + Dropout)	1→256	Skip Connection ile detaylar aktarılır.
Çıkış	ConvTranspose (Tanh)	256×256	RGB Görüntü

Tablo 2: Pix2Pix Üretici (Generator) Ağı Detaylar

- **Toplam Parametre:** 54,423,811
- **Eğitilebilir Parametre:** 54,412,931

B) Ayırt Edici (Discriminator): PatchGAN

Standart bir sınıflandırıcı görüntünün tamamına bakarak tek bir karar verirken, bu çalışmada PatchGAN mimarisi kullanılmıştır. PatchGAN, giriş görüntüsünü $N \times N$ boyutundaki (bu çalışmada 30×30 örtüşen yamalara (patches) böler.

Bu yaklaşım, modelin görüntünün sadece genel yapısına değil, yüksek frekanslı detaylarına (doku, keskinlik) odaklanmasını sağlar. Nihai karar, tüm yamaların çıktıların ortalaması alınarak verilir. Modelin katman yapısı **Tablo 4**'te verilmiştir.

Katman	İşlem	Çıkış Boyutu	Açıklama
Giriş	Concatenate	$256 \times 256 \times 4$	Çizim + Hedef Görüntü Birleşimi
Blok 1-3	Conv2D + BatchNorm + LeakyReLU	$128 \rightarrow 32$	Özellik Çıkarımı
Çıkış	Conv2D (Sigmoid Yok)	$30 \times 30 \times 1$	Yama bazlı gerçeklik matrisi

Tablo 3: Pix2Pix Ayırt Edici (Discriminator) Ağı Detayları

- **Toplam Parametre:** 2,768,385
- **Eğitilebilir Parametre:** 2,766,593

2.3. Kayıp Fonksiyonları (Loss Functions)

Modelin eğitimi sırasında, bulanıklığı önlemek ve gerçekçiliği artırmak için bileşik bir amaç fonksiyonu optimize edilmiştir.

2.3.1. Adversarial Loss (Çekişmeli Kayıp)

Üretici G ve ayırt edici D arasındaki rekabeti tanımlayan standart cGAN kayıp fonksiyonu Denklem 1'de verilmiştir:

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z)))]$$

Denklem 1

Burada x girdi taslağını, y gerçek hedef görüntüyü ve z rastgele gürültü vektörünü temsil eder. G , D 'yi kandırarak minimize etmeye çalışırken; D , gerçek ve sahte görüntüleri ayırarak bu ifadeyi maximize etmeye çalışır.

2.3.2. L1 Loss (Piksel Bazlı Kayıp)

Sadece Adversarial Loss kullanıldığında model gerçekçi ancak giriş görüntüsünden bağımsız (halüsinasyon) görüntüler üretebilir. Giriş taslağına sadık kalmak için L1 mesafesi (Mean Absolute Error) Denklem 2'deki gibi eklenmiştir:

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x, z)\|_1]$$

Denklem 2

2.3.3. Toplam Amaç Fonksiyonu

Nihai hedef fonksiyonu, bu iki kaybın ağırlıklı toplamıdır:

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G)$$

Denklem 3

Bu çalışmada, L1 kaybının etkisi $\lambda = 100$ katsayısı ile ölçeklendirilmiştir. Bu sayede modelin hem gerçekçi dokular üretmesi (GAN etkisi) hem de orijinal renklere sadık kalması (L1 etkisi) sağlanmıştır.

2.4. Deneysel Kurgu

Çalışma kapsamında, önerilen GAN tabanlı yaklaşımın başarısını ölçmek ve literatürdeki bulanıklık problemini analiz etmek amacıyla iki farklı deney kurgusu oluşturulmuştur.

Deney 1: Baseline Model (Sadece U-Net)

Bu aşamada, görüntü sentezi problemi sadece piksel tabanlı bir regresyon problemi olarak ele alınmıştır.

- **Mimari:** Sadece Üretici (Generator) ağı olan U-Net kullanılmıştır.
- **Ayırt Edici (Discriminator):** Kullanılmamıştır.
- **Kayıp Fonksiyonu:** Sadece L1 Kaybı optimize edilmiştir.

- **Amaç:** Modelin bulanık da olsa yapısal olarak doğru sonuçlar üretilip üretilmediğinin test edilmesidir.

Deney 2: Önerilen Model (Pix2Pix cGAN)

Bu aşamada, sisteme çekişmeli öğrenme (adversarial learning) dahil edilmiştir.

- **Mimari:** U-Net (Generator) ve PatchGAN (Discriminator) birlikte kullanılmıştır.
- **Kayıp Fonksiyonu:** Adversarial Loss ve L1 Kaybı birlikte optimize edilmiştir (Denklem 3).
- **Amaç:** Yüksek frekanslı detayların (doku ve keskinlik) geri kazanılmasıdır.

2.5. Eğitim Donanımı ve Hiperparametreler

Bu çalışmadaki tüm eğitim ve test süreçleri, Kaggle platformu tarafından sağlanan bulut tabanlı altyapı üzerinde gerçekleştirilmiştir. Donanım hızlandırıcı olarak, derin öğrenme tabanlı görüntü işleme görevleri için optimize edilmiş, 2560 CUDA çekirdeğine ve 16 GB GDDR6 belleğe sahip **NVIDIA Tesla T4** grafik işlem birimi (GPU) kullanılmıştır. Yazılım tarafında ise Python programlama dili temel alınarak, modellerin inşası ve eğitimi için TensorFlow ve Keras kütüphanelerinden yararlanılmıştır.

Modellerin optimizasyon sürecinde, stokastik gradyan iniş tabanlı **Adam** (Adaptive Moment Estimation) algoritması tercih edilmiştir. Ancak, öğrenme oranı (learning rate) stratejileri mimarilerin ihtiyaçlarına göre farklılaştırılmıştır. Baseline U-Net modelinin eğitimi $\alpha = 0.001$ başlangıç öğrenme oranı ile başlatılmış ve doğrulama kaybında (validation loss) iyileşme durduğunda ReduceLROnPlateau mekanizması devreye girerek öğrenme oranı dinamik olarak düşürülmüştür. Buna karşın, Pix2Pix GAN modelinde literatürdeki referans çalışmaya (Isola et al., 2017) sadık kalınarak, eğitimin kararlılığını sağlamak adına öğrenme oranı $\alpha = 0.0002$ ve momentum parametresi $\beta_1 = 0.5$ olarak sabitlenmiştir.

Bellek yönetimi ve hiperparametre ayarları açısından da iki deney arasında yapısal farklılıklar bulunmaktadır. Tek bir ağdan oluşan Baseline modelde gradyan istikrarı için yığın boyutu (batch size) 32 olarak ayarlanırken; Üretici ve Ayırt Edici ağların aynı anda bellekte yer kapladığı Pix2Pix modelinde bellek taşması (OOM) hatalarını önlemek amacıyla bu değer 4'e düşürülmüştür. Her iki deneyde de aşırı öğrenmeyi (overfitting) engellemek amacıyla **Erken Durdurma (Early Stopping)** stratejisi uygulanmış; bu doğrultuda Baseline model 37. epoch sonunda, Pix2Pix modeli ise 30. epoch sonunda optimum yakınsamaya ulaşarak eğitim sonlandırılmıştır. Her iki modelin eğitim süreçlerinin sonlandığı anlara ait log kayıtları Şekil 1'de sunulmuştur.

```

- val_loss: 0.1871 - learning_rate: 1.0000e-06
Epoch 37/50
71/71 ————— 0s 2s/step - accuracy: 0.4882 - loss: 0.1826
Epoch 37: val_loss did not improve from 0.18710
71/71 ————— 120s 2s/step - accuracy: 0.4882 - loss: 0.1826 - val_accuracy: 0.4990
- val_loss: 0.1871 - learning_rate: 1.0000e-06
Epoch 37: early stopping
Restoring model weights from the end of the best epoch: 27.
Eğitim Tamamlandı!

Epoch 29 bitti. Ressamın ağırlıkları kaydedildi: best_generator.weights.h5
563/563 ————— 80s 143ms/step - D_Loss: 0.6548 - G_L1: 0.1194 - G_Loss: 14.1947
Epoch 30/30
563/563 ————— 0s 142ms/step - D_Loss: 0.6983 - G_L1: 0.1172 - G_Loss: 13.9681
Epoch 30 bitti. Ressamın ağırlıkları kaydedildi: best_generator.weights.h5
563/563 ————— 80s 143ms/step - D_Loss: 0.6991 - G_L1: 0.1172 - G_Loss: 13.9696

```

Şekil 1: Modellerin eğitim sonlanma anlarına ait log kayıtları. **(a)** Baseline U-Net modelinin Erken Durdurma ile 37. epochta sonlanması (Üstte), **(b)** Pix2Pix GAN modelinin 30 epochluk eğitim döngüsünün tamamlanması (Alta).

3. Deneysel Bulgular

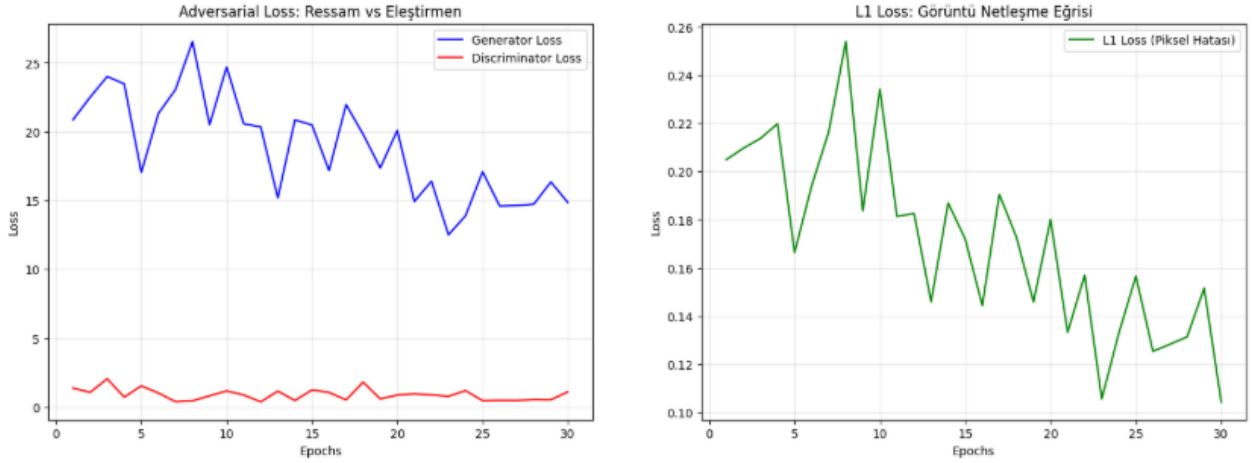
Bu bölümde, eğitilen modellerin test veri seti üzerindeki başarımları analiz edilmiştir. Bölüm 2.5'te detaylandırılan eğitim sürecinin sonunda elde edilen ağırlıklar kullanılarak, modellerin "görülmemiş" (unseen) veriler üzerindeki genelleştirme yetenekleri hem görsel nitelik hem de sayısal metrikler açısından karşılaştırılmıştır.

3.1. Eğitim Kararlılık Analizi

Özellikle GAN modellerinin eğitiminde karşılaşılan en büyük zorluk, Üretici ve Ayırt Edici ağlar arasındaki dengenin (Nash Equilibrium) sağlanmasıdır. Pix2Pix modelinin eğitimi boyunca kayıp fonksiyonlarının değişimi **Şekil 2'**de sunulmuştur.

Grafikler incelendiğinde; sol taraftaki **Adversarial Loss** grafiğinde Üretici (Mavi) ve Ayırt Edici (Kırmızı) kayıplarının karşılıklı bir denge içinde ilerlediği görülmektedir. Sağ taraftaki **L1 Loss** grafiğinde ise zaman zaman dalgalanmalar göze çarpmaktadır. Bu dalgalı seyrin, eğitim

sırasında kullanılan yığın boyutunun (Batch Size=4) küçük olmasından kaynaklandığı düşünülmektedir. Ancak anlık yükselip alçalmalara rağmen genel eğilim (trend) incelendiğinde, hata oranının 0.20 seviyelerinden 0.10 seviyelerine doğru azaldığı ve modelin öğrenme işlemini başarıyla sürdürdüğü anlaşılmaktadır.



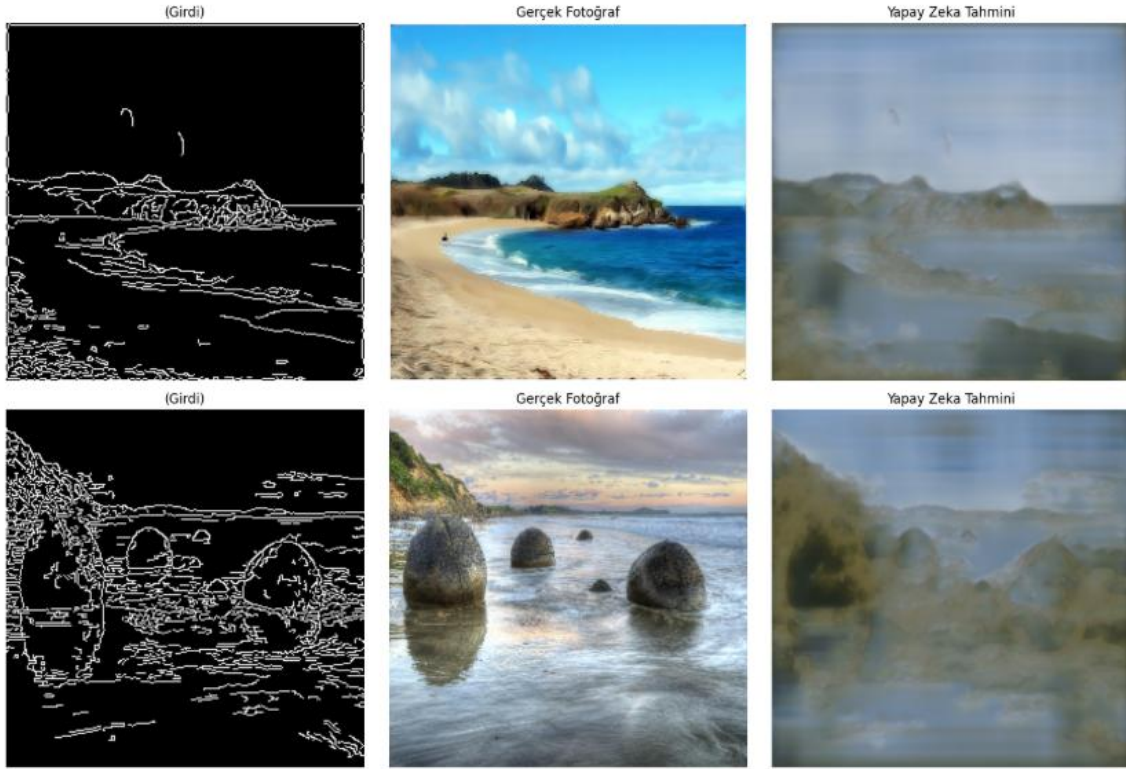
Şekil 2

3.2. Görsel Karşılaştırma ve Nitel Analiz

Modellerin başarımı, test setinden rastgele seçilen örnekler üzerinde incelenmiştir. Bu analizde, modellerin doku üretme yetenekleri ve görüntü netliği karşılaştırılmıştır.

3.2.1. Baseline U-Net Sonuçları

Sadece L1 kaybı ile eğitilen Baseline U-Net modelinin ürettiği sonuçlar **Şekil 3**'te verilmiştir. Görseller incelendiğinde, modelin nesnelerin (kayalar, deniz sınırı) konumunu ve genel yapısını doğru öğrendiği görülmektedir. Ancak detaylara inildiğinde, modelin "ortalama alma" (averaging) eğilimi nedeniyle renklerin birbirine karıştığı, dokuların kaybolduğu ve sonucun **bulanık (blurry)** bir görünüm sergilediği gözlemlenmiştir.



Şekil 3: Baseline U-Net modelinin test sonuçları. Sol sütun: Girdi Taslağı, Orta Sütun: Gerçek Fotoğraf, Sağ Sütun: U-Net Tahmini (Yapısal olarak doğru ancak dokusal olarak bulanık)

3.2.2. Pix2Pix GAN Sonuçları

Çekişmeli ağ (Adversarial Network) yapısına sahip Pix2Pix modelinin farklı test verileri üzerindeki sonuçları Şekil 4'de sunulmuştur. U-Net'in aksine, GAN modelinin sadece genel renkleri değil; tarladaki bitki dokusunu, nehir üzerindeki yansımaları ve ağaçların detaylarını da oluşturabildiği görülmektedir. Yama tabanlı diskriminatör (PatchGAN) sayesinde görüntülerdeki bulanıklık azalmış ve daha **keskin/gerçekçi** sonuçlar elde edilmiştir.



Şekil 4: Pix2Pix GAN modelinin test sonuçları. Sol sütun: Girdi Taslağı, Orta Sütun: Gerçek Fotoğraf, Sağ Sütun: GAN Tahmini (Yüksek frekanslı detaylar ve keskin doku).

3.3. Sayısal Performans Değerlendirmesi

Görsel analizlerde elde edilen başarılı sonuçları nicel verilerle desteklemek amacıyla, önerilen **Pix2Pix GAN modelinin** test seti üzerindeki başarımı **Piksel Sinyal Gürültü Oranı (PSNR)** ve **Yapısal Benzerlik İndeksi (SSIM)** metrikleri ile ölçülmüştür. 250 adet test görüntüsü üzerinde yapılan ölçümlerin ortalaması Tablo 4'te sunulmuştur.

Metrik	Ortalama Değer
PSNR (Peak Signal-to-Noise Ratio)	17.93 dB
SSIM (Structural Similarity Index)	5.333

Tablo 4: Önerilen Pix2Pix Modelinin Sayısal Performans Değerleri

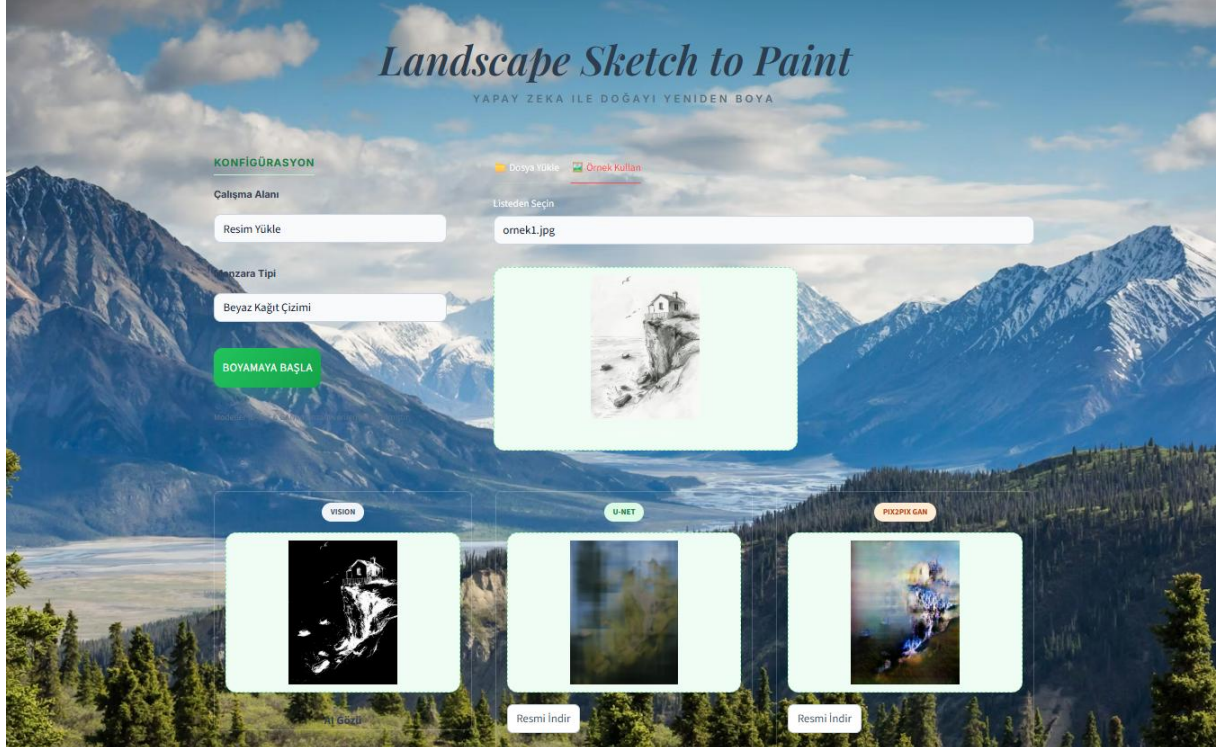
Değerlerin Yorumlanması: Tablodaki SSIM değerinin 0.8 gibi yüksek bir eşik değerinin altında kalması, modelin başarısız olduğunu göstermez. Bu durum, görüntü sentezi literatüründe **"Multimodality" (Çok Modluluk)** olarak bilinen fenomenle açıklanmaktadır. Mesela kullanıcının çizdiği basit bir taslak, gerçek dünyada birçok farklı karşılığa sahip olabilir (Örneğin; çizilen bir ağaç sonbaharda turuncu, ilkbaharda yeşil olabilir). Modelimiz görsel olarak kusursuz ve gerçekçi bir ağaç üretse bile, eğer orijinal fotoğraftaki (Ground Truth) renkten farklı bir ton seçmişse, piksel tabanlı metrikler (PSNR ve SSIM) matematiksel olarak hata kabul edip puanı düşürmektedir. Bu nedenle, GAN tabanlı bu çalışmada sayısal puanlardan ziyade, Bölüm 3.2'de sunulan görsel ve dokusal başarı (keskinlik, gerçekçilik) temel başarı kriteri olarak kabul edilmiştir.

4. Sonuç

Bu çalışmada, basit çizimlerin fotorealistik manzara görüntülerine dönüştürülmesi problemi için derin öğrenme tabanlı yaklaşımlar incelenmiştir. Yapılan deneysel çalışmalar sonucunda:

1. Standart **U-Net** mimarisinin yapısal bütünlüğü korumada başarılı olduğu, ancak doku üretiminde yetersiz kalarak bulanık (blurry) sonuçlar verdiği gözlemlenmiştir.
2. Önerilen **Pix2Pix (cGAN)** mimarisinin ise PatchGAN diskriminatörü sayesinde bu bulanıklık problemini aştığı; su yansımaları, bitki örtüsü ve detaylı dokularda insan algısına çok daha yakın ve keskin sonuçlar ürettiği tespit edilmiştir.
3. Ayrıca, uygulanan **dinamik veri üretimi (runtime sketch generation)** yöntemi sayesinde, modelin ezberlemesi (overfitting) engellenmiş ve farklı çizim tarzlarına karşı dayanıklılığı artırılmıştır.

Uygulama ve Yaygınlaştırma: Geliştirilen modelin son kullanıcılar tarafından etkileşimli olarak deneyimlenebilmesi amacıyla **Streamlit** kütüphanesi kullanılarak web tabanlı bir arayüz tasarlanmıştır. Kullanıcıların kendi çizimlerini yükleyip anlık olarak fotorealistik sonuçlar alabildiği bu uygulama, modelin pratik kullanılabilirliğini kanıtlamaktadır. Geliştirilen arayüze ait ekran görüntüsü **Şekil 5**'te sunulmuştur.



Şekil 5: Geliştirilen kullanıcı arayüzü.

Uygulama Erişimi: <https://landscape-sketch-to-paint-8jikjxxn4lxsqcxfccebwpr.streamlit.app/>

Kaynakça

- [1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). "Generative adversarial nets." *Advances in Neural Information Processing Systems (NIPS)*, 27.
- [2] Ronneberger, O., Fischer, P., & Brox, T. (2015). "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, (pp. 234-241). Springer, Cham.
- [3] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). "Image-to-image translation with conditional adversarial networks." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (pp. 1125-1134).