

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ÜRETİCİ YAPAY NÖRON AĞLARININ MİMARİ TASARIM İÇİN KULLANILMASI

Mukaddes UÇAR

YÜKSEK LİSANS TEZİ
Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Programı

Danışman
Doç. Dr. Mine Elif KARSLIGİL

Temmuz, 2020

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ÜRETİCİ YAPAY NÖRON AĞLARININ MİMARİ TASARIM İÇİN
KULLANILMASI

Mukaddes UÇAR tarafından hazırlanan tez çalışması 27.07.2020 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Programı **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Doç. Dr. Mine Elif KARSLIGİL
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Doç. Dr. Mine Elif KARSLIGİL, Danışman
Yıldız Teknik Üniversitesi

Prof. Dr. Yusuf Sinan AKGÜL, Üye
Gebze Teknik Üniversitesi

Doç. Dr. Gökhan BİLGİN, Üye
Yıldız Teknik Üniversitesi

Danışmanım Doç. Dr. Mine Elif KARSLIGİL sorumluluğunda tarafımca hazırlanan Üretici Yapay Nöron Ağlarının Mimari Tasarım için Kullanılması başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Mukaddes UÇAR

İmza

*Sevgili aileme
ve
biricik eşime*



TEŞEKKÜR

Lisansüssü eğitimim esnasında ders ve tez çalışması dönemlerinde karşılaştığım zorluklarda akademik ve motivasyon desteğini esirgemeyen danışmanım Doç. Dr. Mine Elif KARSLIGİL'e teşekkürlerimi sunarım. Beni yetiştiren ve bugüne kadar verdiğim kararlarda daima arkamda duran ve sonsuz sevgileriyle ve dualarıyla beni dünyanın en değerli varlığı olarak hissettiren aileme minnettarım. Bana, bilgisayar mühendisliğinin sabır gerektiren bir meslek olduğunu öğreten ve teknik ve manevi desteğiyle her zaman yanımda duran sevgili eşim Davut UÇAR'a teşekkür ederim. Akademik başarımın ve çalışma azmimin kaynağı olan orta okul matematik öğretmenim Gökhan YALABIK'ı hayatım boyunca bir gülümseme ve minnetle hatırlayacağım.

Mukaddes UÇAR

İÇİNDEKİLER

SİMGE LİSTESİ	vii
KISALTMA LİSTESİ	viii
ŞEKİL LİSTESİ	ix
TABLO LİSTESİ	xiii
ÖZET	xiv
ABSTRACT	xvi
1 GİRİŞ	1
1.1 Literatür Özeti	1
1.2 Tezin Amacı	3
1.3 Orijinal Katkı	3
2 YÖNTEM	5
2.1 Veri Kümesinin Hazırlanması	5
2.2 Model Mimarisi	7
2.3 Min-Maks Oyunu ve Kayıp Fonksiyonu	11
2.4 Kayıp Fonksiyonu	13
2.5 Optimizasyon Problemi	14
2.6 Regülerizasyon Teknikleri	16
2.6.1 Normalizasyon	16
2.6.2 İki Zamanlı Güncelleme Kuralı	18
2.7 Performans Değerlendirme Metrikleri	18
3 DENEYSEL SONUÇLAR	19
3.1 Değerlendirme Metrikleri ile Deneysel Sonuçların İncelenmesi	20
3.2 Üretilen Yapı Örneklerinin Mimari Açından Değerlendirilmesi	21
3.2.1 Romanesk Mimarisi Örneklerinin Mimari Açından Değerlendirilmesi	21
3.2.2 İslam Mimarisi Örneklerinin Mimari Açından Değerlendirilmesi	23

3.3	Üretilen Yapı Örneklerinin Uzman Tarafından Değerlendirilmesi	26
3.3.1	Üretilen Örneklerde Başarısız Özelliklerin Değerlendirilmesi . .	27
3.4	Üretilen Görüntülerin Gerçekçiliği Üzerine Anket Çalışması	31
3.5	Modelin Ezberleme Probleminin Değerlendirilmesi	38
4	SONUÇ VE ÖNERİLER	46
	KAYNAKÇA	50
A	Derin Öğrenme Teknikleriyle Üretilen Görüntülerin Gerçekçiliğini Ölçme Anketi	53
B	Gerçek ve Üretilen Görüntü Veri Kümesi Örnekleri	68
	TEZDEN ÜRETİLMİŞ YAYINLAR	78

SİMGE LİSTESİ

W	Ağ Ağırlık Parametreleri
D	Ayırt Edici Ağ
$\mathbb{E}$	Beklenen Değer
L	Kayıp Fonksiyonu
N	Normal Dağılım
U	Üniform Dağılım
G	Üretici Ağ

KISALTMA LİSTESİ

AdaGrad	Adaptive Gradient
BigGAN	Big Generative Adversarial Networks
Bknz	Bakınız
BS	Başlangıç Skoru
CMYK	Cyan-Magenta-Yellow-Key
FBM	Fréchet Başlangıç Mesafesi
FID	Fréchet Inception Distance
GAN	Generative Adversarial Networks
GN	Grup Normalizasyonu
KL-Uzaklığı	Kullback-Leibler Uzaklığı
KN	Küme Normalizasyonu
Min-Maks	Minimum-Maksimum
ResNet	Residual Networks
RGB	Red-Green-Blue
RMSProp	Root Mean Square Propagation
SAGAN	Self-Attention Generative Adversarial Networks
TTUR	Two-Timescale Update Rule
ÜÇA	Üretici Çekişmeli Ağlar
VOK	Varyasyonel Otokodlayıcılar
YNA	Yapay Nöron Ağları

ŞEKİL LİSTESİ

Şekil 2.1	Görüntü veri kümesi Romanesk mimarisi örnekleri	6
Şekil 2.2	Görüntü veri kümesi İslam mimarisi örnekleri	7
Şekil 2.3	ResNet blok yapısı	8
Şekil 2.4	ResNet geri besleme modülü	8
Şekil 2.5	ResNet blok aşamaları	9
Şekil 2.6	Üretici (solda) ve ayırt edici (sağda) ağın mimari şablonları	10
Şekil 2.7	Öz ilgi modülünün işleyişi	11
Şekil 2.8	Üretici çekişmeli ağ mimarisi	12
Şekil 2.9	Grup normalizasyonu	17
Şekil 3.1	Eğitilen modelin Fréchet Başlangıç Mesafesi ve Başlangıç Skoru sonuçları (sırasıyla)	20
Şekil 3.2	Romanesk mimarisine ait gerçek (üstte) ve üretilen (altta) dış cephe örnekleri	21
Şekil 3.3	Romanesk mimarisine ait gerçek (üstte) ve üretilen (altta) tonoz ve sunak örnekleri	22
Şekil 3.4	Romanesk mimarisine ait gerçek (üstte) ve üretilen (altta) çan kulesi ve kapı girişi örnekleri	23
Şekil 3.5	Romanesk mimarisine ait gerçek (üstte) ve üretilen (altta) vitray ve kapı süslemesi örnekleri	23
Şekil 3.6	İslam mimarisine ait gerçek (üstte) ve üretilen (altta) dış cephe örnekleri	24
Şekil 3.7	İslam mimarisine ait gerçek (üstte) ve üretilen (altta) eyvan örnekleri	25
Şekil 3.8	İslam mimarisine ait gerçek (üstte) ve üretilen (altta) minare ve kubbe örnekleri	25
Şekil 3.9	İslam mimarisine ait gerçek (üstte) ve üretilen (altta) mukarnas ve süsleme örnekleri	26
Şekil 3.10	Üretilen İslam mimarisi (üstte) ve Romanesk mimarisi (altta) görüntü örnekleri	27
Şekil 3.11	Romanesk mimarisi gerçek görüntü örnekleri	27
Şekil 3.12	İslam mimarisi gerçek görüntü örnekleri	28
Şekil 3.13	Üretilen hatalı kubbe örnekleri	28

Şekil 3.14 Üretilen hatalı kubbe örnekleri	29
Şekil 3.15 Üretilen hatalı minare örnekleri	29
Şekil 3.16 Üretilen hatalı İslam mimarisi örnekleri	29
Şekil 3.17 Üretilen Romanesk mimarisi örnekleri	30
Şekil 3.18 Romanesk mimari üretilen dış cephe örnekleri	30
Şekil 3.19 Romanesk mimarisi üretilen taç kapı girişi örnekleri	31
Şekil 3.20 Üretilen romanesk mimari kemer örnekleri	31
Şekil 3.21 14 sorudan oluşan ankette katılımcıların görüntülere verdiği gerçekçilik puanları	32
Şekil 3.22 Anket çalışmasında gerçekçi olarak ifade edilen örneklerden bazıları (Görüntü çiftlerinde üretilen görüntü solda ve gerçek görüntü sağda yer almaktadır)	33
Şekil 3.23 Katılımcılar tarafından gerçekçi bulunmayan görüntüler (Görüntü çiftlerinde üretilen görüntü solda ve gerçek görüntü sağda yer almaktadır)	34
Şekil 3.24 14 sorudan oluşan ankette katılımcıların üretilen ve gerçek görüntülere verdiği gerçekçilik puanları	35
Şekil 3.25 Katılımcıların (sırasıyla) 54%, 61%, 73% ve 49%'u tarafından gerçekçi veya gerçeğe yakın olarak oylanan İslam mimarisi örnekleri	36
Şekil 3.26 Katılımcıların (sırasıyla) 73%, 78%, 83% ve 88%'i tarafından gerçekçi veya gerçeğe yakın olarak oylanan İslam mimarisi örnekleri	37
Şekil 3.27 14 sorudan oluşan ankette mimaride uzman katılımcıların üretilen ve gerçek görüntülere verdiği gerçekçilik puanları	38
Şekil 3.28 Ezberleme problemini tespit etmede uygulanan adımlar	40
Şekil 3.29 Üretilen (a), en benzer gerçek (b) görüntü, kırpma-yeniden boyutlandırma (c), rasgele karartma (d, e, f, g, h), Laplacian filtresi (i), Gaussian gürültüsü (j) veya tuz-biber gürültüsü (k, l, m, n) uygulanmış görüntüler ve orijinal gerçek görüntüyle benzerlik mesafeleri	40
Şekil 3.30 Üretilen (a), en benzer gerçek (b) görüntü, kırpma-yeniden boyutlandırma (c), rasgele karartma (d, e, f, g, h), Laplacian filtresi (i), Gaussian gürültüsü (j) veya tuz-biber gürültüsü (k, l, m, n) uygulanmış görüntüler ve orijinal gerçek görüntüyle benzerlik mesafeleri	41
Şekil 3.31 Üretilen örneklerle en benzer gerçek örneklerin Öklid mesafeleri(üstte), çeşitli gürültüler uygulanmış gerçek örneklerin orijinal örneklerle minimum Öklid mesafeleri(alta)	41

Şekil 3.32	Üretilen görüntülerin gerçek görüntülerle benzerlik mesafeleri (üstte) ve rasgele kırılmış, Gaussian gürültüsü eklenmiş ve Laplacian filtresi uygulanan görüntülerin gerçek görüntülerle benzerlik mesafeleri (altta)	42
Şekil 3.33	Üretilen görüntülerin gerçek görüntülerle benzerlik mesafeleri (üstte) ve 1/25'i, 1/13'ü ve 1/8'i rasgele karartılan görüntülerin orijinal gerçek görüntülerle benzerlik mesafeleri (altta)	43
Şekil 3.34	Üretilen görüntülerin gerçek görüntülerle benzerlik mesafeleri (üstte) ve 1/5'i ve 1/4'ü rasgele karartılan görüntülerin orijinal gerçek görüntülerle benzerlik mesafeleri (altta)	44
Şekil 3.35	Üretilen görüntülerin gerçek görüntülerle benzerlik mesafeleri (üstte) ve farklı oranlarda tuz-biber gürültüsü eklenmiş görüntülerin orijinal gerçek görüntülerle benzerlik mesafeleri (altta)	44
Şekil 3.36	Üretilen veya transformasyon uygulanmış görüntülerin orijinal gerçek görüntülerle ortalama mesafesi ve sapma değerleri	45
Şekil A.1	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 1. soru	54
Şekil A.2	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 2. soru	55
Şekil A.3	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 3. soru	56
Şekil A.4	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 4. soru	57
Şekil A.5	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 5. soru	58
Şekil A.6	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 6. soru	59
Şekil A.7	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 7. soru	60
Şekil A.8	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 8. soru	61
Şekil A.9	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 9. soru	62
Şekil A.10	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 10. soru	63
Şekil A.11	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 11. soru	64
Şekil A.12	Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 12. soru	65

Şekil A.13 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 13. soru	66
Şekil A.14 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 14. soru	67
Şekil B.1 Veri kümesinde yer alan Romanesk mimarisine ait genel dış cephe örnekleri	68
Şekil B.2 Romanesk mimarisine ait veri kümesinde yer alan başlıca unsurlar .	69
Şekil B.3 Veri kümesinde yer alan İslam mimarisine ait genel dış cephe örnekleri	70
Şekil B.4 İslam mimarisine ait veri kümesinde yer alan başlıca unsurlar	71
Şekil B.5 Model tarafından üretilmiş Romanesk mimari dış cephe örnekleri . .	72
Şekil B.6 Romanesk mimarisine ait üretilmiş görüntülerde yer alan unsurlar .	73
Şekil B.7 Romanesk mimarisine ait üretilmiş görüntülerde yer alan unsurlar .	74
Şekil B.8 Model tarafından üretilmiş İslam mimarisi dış cephe örnekleri	75
Şekil B.9 İslam mimarisine ait üretilmiş görüntülerde yer alan unsurlar	76
Şekil B.10 İslam mimarisine ait üretilmiş görüntülerde yer alan unsurlar	77

TABLO LİSTESİ

Tablo 2.1	Üretici ve ayırt edici ağın hedef fonksiyonları	12
------------------	---	----



Üretici Yapay Nöron Ağlarının Mimari Tasarım için Kullanılması

Mukaddes UÇAR

Bilgisayar Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Danışman: Doç. Dr. Mine Elif KARSLIGİL

Üretici Çekişmeli Ağlar (ÜÇA) birbiriyle rekabet eden iki ayrı derin yapay nöron ağından oluşan, yeni ve gerçekçi yapay örnekler üretmeyi amaçlayan üretici modellerdir. Bu iki ağdan birisi olan üretici ağ rassal gürültü vektörünü girdi olarak alır ve gürültüyü yapay bir görüntü örneğine dönüştürerek çıktı olarak verir. Üretici Çekişmeli Ağ'ı oluşturan diğer ağ ise ayırt edici ağıdır. Ayırt edici ağ, üretici ağdan elde edilen görüntü ile gerçek bir görüntü örneğini girdi olarak alır ve her birini "üretilmiş görüntü" veya "gerçek görüntü" olarak sınıflandırır. Ağların optimizasyonu için eğitim aşamasında her iki ağın kayıp fonksiyonları farklı biçimde hesaplanır. Üretici ağda üretilmiş örnek için kayıp hesaplanarak parametrelerin optimizasyonu yapılır. Ayırt edici ağda ise hem gerçek hem de üretilmiş örnekler için toplam kayıp hesaplanarak ağ parametrelerine geri besleme uygulanır. Birbiriyle rekabet eden iki ağdan birisinin parametre güncellemesi diğer ağın parametrelerini değişime zorlamıyorsa parametre eğitimi sonlandırılır. Üretici ağın güncel parametre değerleriyle yapay örnekler üretilir.

Eldeki veri kümesinin olasılıksal dağılımını öğrenerek, bu dağılıma uygun örnekler üretmeyi amaçlayan ÜÇA ile yapay görüntü ve video üretme konusunda çeşitli çalışmalar yapılmıştır. Bu çalışma ile derin ağların üretici modellerinden birisi olan Üretici Çekişmeli Ağlar'ın farklı mimari tarzlarda bina ve bina unsurlarını üretmedeki başarımları üzerine bir araştırma yapılmıştır. ÜÇA yardımıyla Romanesk ve İslam mimarisi gibi farklı mimari tarzlarda gerçekçi yapılar üretilmiştir. Eğitimde kullanılan görüntü veri kümesi yaklaşık 39000 cami, kilise, katedral, medrese, kütüphane ve kalelere ait görüntülerden oluşturulmuştur. Veri kümesinde binaların dış ve iç cephe

görüntülerinin yanında kubbe, kule, minare, tonoz, kapı girişi ve vitray gibi mimari unsurlar bulunmaktadır. Bu çalışmada Üretici Çekişmeli Ağlar ile gerçekçi ve yüksek kalitede Romanesk ve İslam mimarisi örnekleri üretmek için BigGAN ağ mimarisi üzerinde gerekli düzenlemeler yapılmıştır.

Üretilen görüntüler ile modelin başarısı ve performans değerlendirme metrikleriyle örnek kalitesi değerlendirilmiştir. Yapılan anket çalışmasıyla görüntülerin gerçekçiliği ve örneklerdeki problemler araştırılmıştır. 48 katılımcının çoğunluğu örneklerin genellikle gerçekçi ya da gerçeğe yakın olduğu görüşündedir. Katılımcılara göre gerçekçi olmayan örneklerde temel problemler mevcuttur: simetri özelliğinin olmaması, detay özelliklerin eksikliği, görüntülerde çizgisel kayma ve kırılmalar ve keskinliğin yetersiz olması. Ayrıca bu anketin bir benzeri mimari alanında uzman kişilere uygulanmıştır. Ankete katılan 24 uzmana göre Romanesk mimarisi örnekleri İslam mimarisi örneklerine göre daha gerçekçidir.

Görüntülerde öne çıkan yapısal özellikler ve modelin üretmekte zorlandığı mimari özellikler incelenmiştir. Ayrıca üretilen örneklerin kalitesi mimaride uzman bir kişi tarafından değerlendirilmiştir. İslam mimarisi örneklerinin çok geniş bir coğrafyada yer alması nedeniyle mimari tarz farklılaşmaktadır. Bu durum da Romanesk mimarisi örnekleriyle karşılaştırıldığında İslam mimarisi örneklerinin daha başarısız olduğunu göstermektedir. Modelin ezberleme yapma olasılığını değerlendirmek için İçerik Tabanlı Görüntü Erişimi yöntemi yardımıyla eğitilmiş ayırt edici ağ kullanılarak yeni bir yöntem geliştirilmiştir. Üretilen görüntülere en benzer gerçek görüntüler tespit edilerek bu görüntülere çeşitli transformasyon işlemleri uygulanmıştır. Üretilen örneklerle karşılaştırıldığında, transformasyon uygulanmış görüntülerin orijinal görüntülere daha benzer olması modelin ezberleme yapmış olma olasılığının düşük olduğunu göstermiştir.

Anahtar Kelimeler: Üretici çekişmeli ağlar, derin öğrenme, mimari tasarım, görüntü üretme

Architectural Design with Generative Adversarial Networks

Mukaddes UÇAR

Department of Computer Engineering

Master of Science Thesis

Advisor: Assoc. Prof. Dr. Mine Elif KARSLIGİL

Generative Adversarial Networks (GANs) are generative models having two distinct non cooperative deep neural networks to generate new realistic image samples. One of the networks is called as a generator which takes random noise vector and transform it into the artificial image sample. Another network is defined as a discriminator which takes real and generated image samples and classifies them by real or fake. The loss functions of two networks are calculated according to the outputs of the discriminator. Gradients of the loss functions are back propagated to the generator and the discriminator to update weight parameters of the networks. Similar steps are implemented to the other image samples to keep go on the training. When the weight updates of a network do not have an effect on behaviour of the other network, training is stopped. The generator creates images using the last updated weight parameters.

There are various studies associated with GANs which aim to represent the distribution of a data set and generate samples fitting the real distribution. In this study, buildings and architecture elements are generated in different styles of architectures using GANs. Realistic samples of Romanesque and Islamic Architecture are generated and success of the model is evaluated. A new data set is created that consists of approximately 39000 mosque, church, cathedral, library, university and castles. The data set involves not only facades but also the elements such as dome, tower, minaret, vault, door entrance and stained glass.

In this study, the BigGAN model is fine-tuned to obtain realistic and high-quality images. Success of the model is assessed from the viewpoint of Performance

Evaluation Metrics.

A survey was arranged to determine that the generated samples whether look like real or not. The majority of 48 respondents think that the samples are generally realistic or nearly realistic. According to the survey fundamental problems of unrealistic samples are lack of the feature of symmetry in images, defects of the detailed features, lack of the straightness and, shaky and unclear shapes of the buildings. Similar of the survey was arranged in architecture experts. They think that samples of the Romanesque architecture were more realistic than samples of the Islamic architecture.

Besides the survey, quality of the generated images was assessed in technique details by an architecture expert. According to the expert generally, samples of the Romanesque style more successful than samples of the Islamic architecture. It is because that samples of the Islamic style have wide variety in point of a building material, a texture and a form. Too much variety causes the under fitting model and poor samples.

There is a possibility of the over fitting of the trained model. To investigate this, a novel method was improved using the method of content-based image retrieval and the trained model. The similarity of the generated and real images was computed. Similarly, the similarity between the transformed real images and the original real images was obtained. The results demonstrated that the transformed images more similar to original real ones than generated images. Eventually, according to the results the possibility of the over fitting is low.

Keywords: Generative adversarial networks, deep learning, architectural design, generating images

1.1 Literatür Özeti

Üretici modeller eldeki verinin dağılım fonksiyonunu temsil etmeyi amaçlar. Yapılan çalışmalar bunu iki yöntemden birisiyle gerçekleştirir: doğrudan veri dağılımını tahmin ederek ya da dağılımı doğrudan tahmin etmeden dağılıma uyan yeni örnekler üreterek. Veri dağılımını doğrudan öğrenerek, dağılıma uyan farklı örnekler üretmek için inanç ağlarından veya otokodlayıcıların farklı bir biçimi olan varyasyonel otokodlayıcılardan yararlanılmıştır.

İnanç ağları (Belief Networks) n-boyutlu bir vektörün olasılıksal dağılımını 1-boyutlu dağılımların çarpımı olarak ifade eden zincir modellerdir [1]. İnanç ağlarıyla ilgili yapılan bir çalışmada [2] görüntünün her bir pikseli ile kendinden önceki pikseller arasındaki ilişki olasılıksal olarak modellenmiş ve görüntünün tamamı için elde edilen olasılıksal dağılım kullanılarak yeni görüntüler elde edilmiştir. Her bir pikselin kendinden önceki piksellere bağlı olarak olasılıksal dağılımının oluşturulması karmaşıklığı büyük oranda artırmıştır. Buna çözüm getirmek için [3] adlı çalışmada her bir pikselin çevresindeki belirli bir bölgedeki piksellerle olan bağımlılığı modellenmeye çalışılmıştır. Bu çalışmada konvolüsyonel filtrelerin kullanılması eğitimi hızlandırırsa da her bir piksel için art arda yapılan olasılık dağılımı hesaplamaları karmaşıklığı azaltmamıştır. Ayrıca hesaplamalar paralel olarak gerçekleştirilememektedir. 2014 yılında otokodlayıcılardan yararlanarak örnek üretimi üzerine bir çalışma yapılmıştır. Otokodlayıcılar, verilen girdiyi kodlayarak daha düşük boyutlarda özellik çıkarımı yapan kodlayıcı ve bu özellikleri kullanarak yeniden girdiyi elde etmeye yarayan çözücüden oluşur. Otokodlayıcılar görüntü sıkıştırma [4–6] ve görüntü renklendirmede [7] başarılı sonuçlar verse de yeni içerik üretemezler. Otokodlayıcılarla örnek üretebilmek için kodlanan girdilerin örtülü uzaydaki özellikleri normal dağılıma uymaya zorlanır. Ardından, normal dağılıma uyan rastgele özellikler üretilir ve bu özelliklerden çözücü yardımıyla yeni görüntüler oluşturulur. Varyasyonel Otokodlayıcılar (VOK) adı verilen bu ağlar kullanılarak görüntü üretme çalışmaları yapılmıştır [8–10]. Bu çalışmalardan [10] VOK'un

doğal eskiz çizimleri üretmede başarılı olduğunu ve yaratıcı taslak resimler çizdiğini göstermiştir. Ancak [8] ve [9] incelendiğinde VOK, şekilleri belirli olan obje, insan veya hayvan figürü gibi görüntüler üretmede yetersiz kalmıştır. Üretilen görüntülerin bulanık olması bu durumun göstergesidir. Üretici modellerde veri dağılımını doğrudan öğrenme veya veri dağılımına uyan yeni örnekler üretme olarak iki farklı yöntem kullanıldığından daha önce bahsedilmiştir. Üretici Çekişmeli Yapay Nöron Ağları (ÜÇA) her iki metotla da model oluşturabilmesine rağmen, yapılan çalışmalar örnek üretimi üzerinde yoğunlaşmıştır [1]. ÜÇA 2014 yılında [11] çalışmasıyla tanıtılmıştır. Bu çalışmada MNIST, CIFAR ve Toronto Face gibi basit veri kümeleri kullanılarak örnekler üretilmiştir. [11] ÜÇA modeli “üretici” ve “ayırt edici” adlı, birbiriyle yarışan çok katmanlı iki perseptron ağından oluşmaktadır. 2016 yılında yapılan [12] çalışmasında ÜÇA, perseptron yerine derin konvolüsyonel ağlar kullanılarak daha kararlı bir model oluşturulmuştur. [13] çalışmasında üretici ve ayırt edici ağ düşük çözünürlükte görüntülerle eğitime başlanmış ve eğitim esnasında genişletilmiş ağ katmanları eklenerek üretilen görüntülerin çözünürlüğü arttırılmıştır. Ancak ImageNet gibi karmaşık veri kümeleri kullanarak daha gerçekçi, yüksek çözünürlüklü ve çok çeşitli görüntüler elde edilmesi 2018 yılına dayanmaktadır. Önceki çalışmalarda ÜÇA objelerin geometrik ve yapısal parçalarını üretmede başarılı değildir. Bu nedenle bir köpek gerçekçi kürk dokusuyla üretilebilse de köpeğin bacakları ayrı ayrı görünmemektedir [14]. [14] çalışmasında oluşturulan model, objelerin parçalarının mekânsal ilişkisini göz önüne alarak görüntü üretmiştir. Objeleri oluşturan parçaların şekli ve uzaysal yerleri bir bütün oluşturmada büyük bir etken olduğu için [14] gerçekçi görüntü üretmede başarılı olmuştur. [14] mimarisini kullanan [15] çalışmasında ise, karmaşık bir veri kümesi olan ImageNet üzerinde başarılı sonuçlar elde edilmiştir. Eğitimdeki girdilerin oluşturduğu grupların büyüklüğünü artırarak, ağı derinleştirerek ve ayırt ediciyi üreticiden daha iki kat fazla eğiterek 128x128, 256x256 ve 512x512 boyutlarında gerçekçi ve çok çeşitli görüntüler üretmeyi başarmıştır.

Bunların yanında, mimari tasarım konusunda yapılan ÜÇA çalışmaları da incelenmiştir. [16]’da bir girdi görüntüsünü kullanarak farklı tarzda görüntü elde etme üzerine bir çalışma yapılmıştır. Kapı, pencere ve balkon gibi bina unsurlarının dış cephede bulunduğu yeri gösteren bir girdiden bina görüntüsü elde edilmiştir. Bu çalışmada rastgele bir gürültüden örnek üretmek yerine girdi olarak verilen görüntü üzerinde dönüşüm işlemi yapılmıştır.

[17] çalışmasında apartman dairesi plan tasarımı yapılmıştır. Model, [16] çalışmasındaki gibi girdi görüntüsünü istenen tarzdaki çıktı görüntüsüne dönüştüren bir üretici ağıdır. Üç adımda ve farklı ağlarla gerçekleştirilen tasarımda diğer Üretici Çekişmeli Ağlar’ın aksine girdi olarak gürültü vektörü yerine dönüşümü yapılacak görüntü alınır. İlk adımda girdi olarak verilen arazi parseline göre binanın kuş bakışı

görünümü ağ tarafından üretilmiştir. İkinci adımda kullanıcı tarafından belirlenen kapı ve pencere yerlerine göre ikinci üretici ağ ile odaların planı oluşturmuştur. Üçüncü ağ ise odalara mobilyaları ekleyerek planı üretir.

[18]'de ÜÇA'nın derin katmanlarını inceleyerek, görüntülerin semantik olarak nasıl üretildiği konusunda bir araştırma yapılmıştır. Gizli katmandaki özellik haritalarında nöronların oluşturduğu görüntü unsurlarını segmentasyonla tespit ederek, bina görüntülerine kapı, pencere ve kemer gibi mimari unsurlar ekleme ve görüntülerden bu unsurları silme üzerine bir çalışma gerçekleştirilmiştir.

1.2 Tezin Amacı

Tarihi 1943'e dayanan derin öğrenme 2000'li yıllardan bugüne endüstriden, hizmet sektörüne, sağlık alanından sosyal ağlara kadar hayatımızın hemen her alanında varlığını sürdürmektedir. Son yıllarda derin öğrenme tabanlı üretici modeller ile işlenmemiş veriyi kullanarak anlamlı bilgi üretme üzerine birçok çalışma yapılmıştır. Görüntü işleme alanında görüntü ve video üretme, görüntü kalitesi iyileştirme, pekiştirmeli öğrenme, veri sıkıştırma ve stil transferi konularında giderek daha başarılı sonuçlar elde edilmektedir.

Üretici modeller yardımıyla çizgi karakter [1], emoji [2], resim [3] ve illüstrasyon [4] çizme gibi yaratıcı sanat konularında da birçok çalışma yapılmıştır. Bu çalışma yukarıda bahsedilen sanat ve tasarım çalışmalarına benzer şekilde, gerçekçi ve yaşanabilir mimari yapıların makineler tarafından tasarlanması için ilk adım projesi niteliğindedir. Proje ile Üretici Çekişmeli Ağlar (ÜÇA) (Generative Adversarial Networks, GAN) kullanarak Romanesk ve İslam mimarisi gibi farklı mimari tarzlarda yapılar üretilmiştir.

1.3 Orijinal Katkı

Bu çalışma ile üretilen görüntülerde öne çıkan yapısal özellikler ve modelin üretmekte zorlandığı mimari özellikler değerlendirilmiştir. Yapılan anket çalışması ile üretilen görüntülerin gerçekçiliği nicel olarak değerlendirilmiş ve örneklerdeki genel problemler tespit edilmiştir.

Görüntülerin kalitesi mimaride uzman anket katılımcıları tarafından değerlendirilmiştir. Uzmanlara göre İslam mimarisine ait örnekler Romanesk mimarisi örneklerine göre daha az başarılıdır. Bu durumun başlıca sebebi İslam mimarisinin Afrika, Asya ve Avrupa kıtalarında çok geniş bir coğrafyada yer almasıdır. Romanesk mimarisi ise Batı Avrupa'da dar bir coğrafyada ve dört yüzyıl gibi kısa bir

zaman dilimini kapsamaktadır. Bu nedenle İslam mimarisi genel bir üslup Romanesk mimarisi ise alt üslup olarak düşünülebilir.

Modelin ezberleme yapmış olması olasılığı eğitilmiş ayırt edici ağ ve içerik tabanlı görüntü erişimi yöntemleri kullanılarak nicel yöntemlerle değerlendirilmiştir. Üretilmiş görüntülerle karşılaştırıldığında, transformasyon işlemi uygulanan gerçek görüntülerin orijinal görüntülerle benzerliği daha fazladır. Bu durum ezberleme olasılığının düşük olduğunu göstermiştir.



2 YÖNTEM

Bu çalışmada derin öğrenme tabanlı üretici modellerden bir olan ÜÇA'nın mimari tasarımıdaki başarımları üzerine bir analiz yapılmıştır. İncelenen model sınıf koşullu görüntü üretme (class-conditional image synthesis) üzerine büyük bir başarı kaydetmiş olan BigGAN [15]'dir. Modelin omurgasını ImageNet üzerinde eğitilmiş olan ResNet50 [19] ve SAGAN [14] modelleri oluşturmaktadır. ÜÇA'nın eğitimi hem model mimarisine hem de optimizasyon parametrelerine karşı çok hassastır. Bu nedenle gerçekçi Romanesk ve İslam mimarisi görüntüleri elde etmek için BigGAN mimarisi üzerinde gerekli düzenlemeler yapılmıştır. Görüntülerin gerçekçiliğini nicel olarak test etmek için 48 katılımcıya sunulan bir anket düzenlenmiştir. Anket sonuçları görüntülerin çoğunluğunun gerçekçi olduğunu göstermiştir. Görüntü kalitesi bir mimari uzmanı tarafından da değerlendirilmiştir. Romanesk mimarisi örnekleri İslam mimarisi örneklerine göre daha başarılı bir biçimde üretilmiştir. Bu durumun sebebi geniş bir coğrafyaya yayılan İslam mimarisi örneklerinin yapı malzemesi, biçim ve dokusal olarak çok çeşitlenmesidir. Karakteristik özelliklere sahip olan alt üslup niteliğindeki Romanesk mimarisini başarılı bir şekilde öğrenen model, aynı başarıyı İslam mimarisinde yakalayamamıştır. Modelin ezberleme olasılığı içerik tabanlı görüntü erişimi yöntemi kullanarak tespit edilmeye çalışılmıştır. Gerçek görüntülere transformasyon uygulayarak bu örneklerin orijinal örneklerle benzerliği ölçülmüştür. Benzer şekilde üretilen örneklerin de gerçek orijinal örneklerle benzerliği hesaplanmıştır. Deneysel sonuçlar transformasyon uygulanan görüntülerin gerçek görüntülere daha benzer olduğunu göstermiştir. Bu durum modelin ezberleme olasılığının az olduğunu göstermiştir.

2.1 Veri Kümesinin Hazırlanması

Bu çalışmada Flickr, Pixabay, Wikipedia ve StockFreeImages paylaşım sayfalarından “#romanesquearchitecture” ve “#islamicarchitecture” hashtag' leri ile herkesin kullanımına açık yaklaşık 20000 İslami mimariye ve yaklaşık 19000 Romanesk mimarisine ait görüntü seçilmiştir. Bu görüntülerin boyutları 129x136 ile 7920x8150

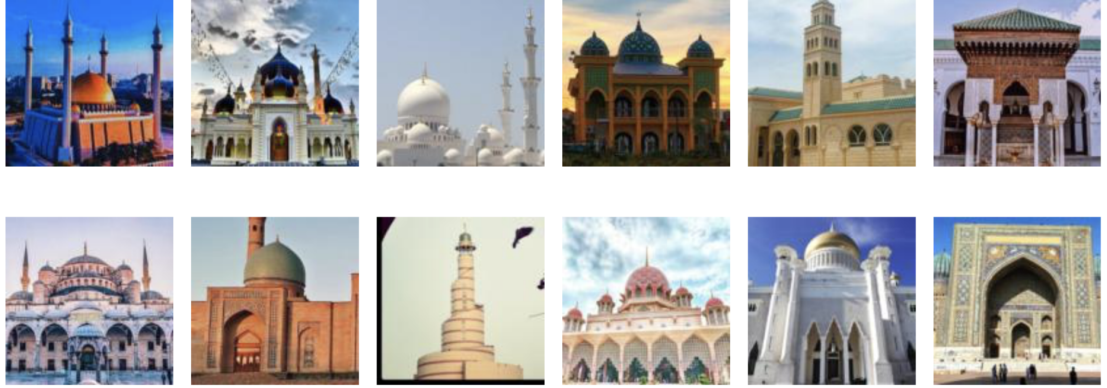
arasında deęişkenlik gösterse de medyanı 857x966 boyutlarıdır. Veri kümesi, cami, kilise, katedral, kütüphane, üniversite ve kalelere ait görüntülerden oluşmaktadır. Ayrıca, binaların dış ve iç cephe görüntülerinin yanında kubbe, kule, minare, tonoz, kapı girişı ve vitray gibi mimari unsurlar bulunmaktadır. Veri kümesini standart ve kullanışlı hale getirmek amacıyla eğitim aşamasından önce görüntüler aşağıda belirtilen önışlemlerden geçirilmiştir:

- RGBA(Red-Green-Blue-Alpha) ve CMYK(Cyan-Magenta-Yellow-Key) renk uzayındaki görüntüler RGB(Red-Green-Blue) 'ye dönüştürülmüştür.
- Veri kümesindeki görüntülerde mimari yapılar sınırlarından kırılarak gereksiz obje ve arkaplan bilgisi azaltılmıştır.
- Derin aęlarla eğitilecek olan görüntülerin standart boyuta getirilmesi ve aęın kompleksliğinden dolayı küçük boyutlarda olması gereklidir. Görüntüler yüksek çözünürlükte olduęu için her bir görüntü 5 kez rastgele kırılarak veri kümesi genişletilmiştir ve boyut küçültmenin sebep olacaęı bilgi kaybı da önlenmeye çalışılmıştır. Böylece 114000 Romanesk ve 120000 İslam Mimarisi görüntüsü elde edilmiştir.
- Yüksek çözünürlüklü görüntüler 128x128 boyutuna düşürülmüştür.



Şekil 2.1 Görüntü veri kümesi Romanesk mimarisi örnekleri

Şekil 2.1 ve 2.2'de veri kümesinden Romanesk ve İslam mimarisine ait örnekler yer almaktadır. İslam mimarisi örnekleriyle karşılaştırıldığında Romanesk mimarisi örnekleri birbirine daha benzerdir. İslam mimarisinde yapıların inşasında kullanılan malzeme ve kubbe, minare ve kemer gibi biçimsel ana unsurların çok çeşitli olması ortak özelliklerin mimari tarz olarak bir bütün oluşturmasını zorlaştırmaktadır. Romanesk mimarisi örnekleri ise hem malzeme hem de biçimsel olarak birbirine



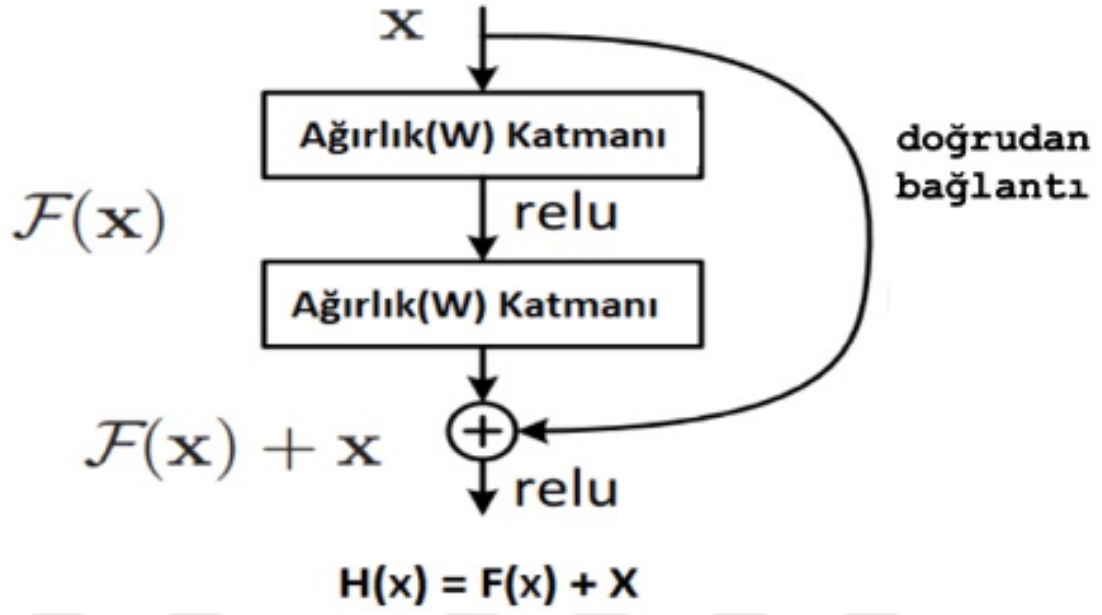
Şekil 2.2 Görüntü veri kümesi İslam mimarisi örnekleri

benzer karakteristik örneklerdir. Bu nedenle İslam mimarisini farklı tip örnekleri kapsayan genel bir üslup, Romanesk mimarisini ise bir alt üslup olarak ele alınabilir. Sonuçta eğitilen modelin her bir mimari tarz için öğrenme ve üretme başarısı farklı olması beklenir. Deneysel Sonuçlar bölümünde bu konuya değinilmiştir. Ayrıca Ekler bölümünde veri kümesinden daha fazla örnek görüntü gösterilmiştir.

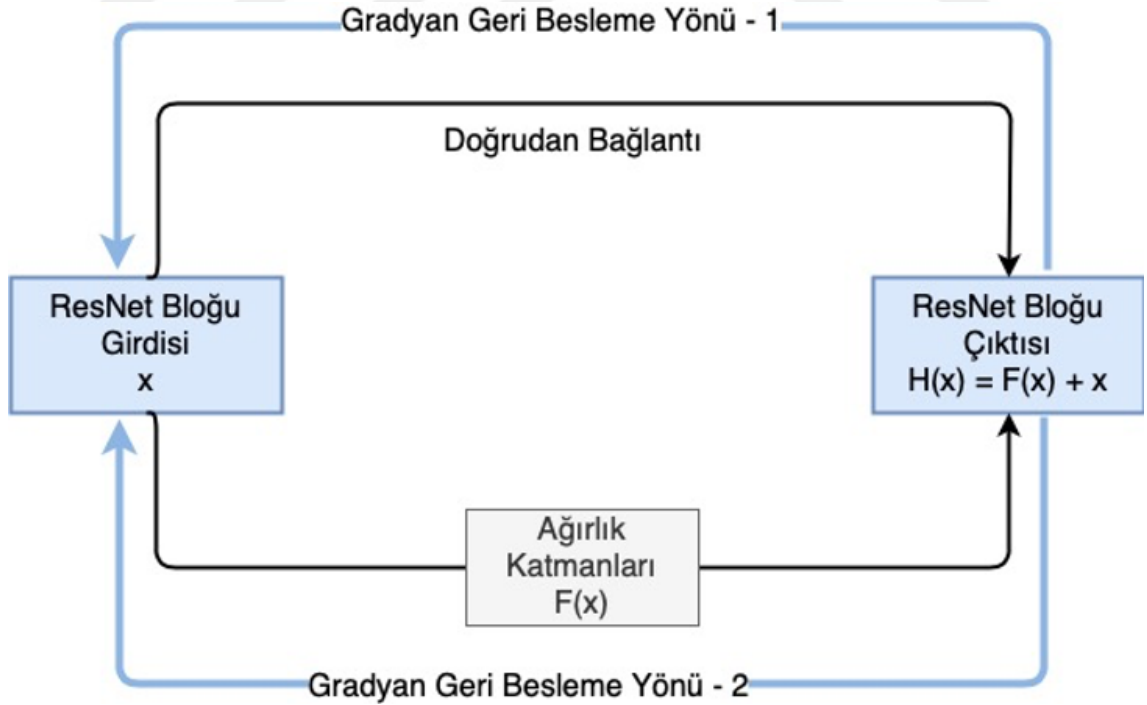
2.2 Model Mimarisi

Bu çalışmada yararlanılan BigGAN modeli ResNet-50 (Residual Networks - 50 Layers) mimarisine dayanmaktadır. Konvolüsyonel ağların makine öğreniminde kullanılması birçok sınıftan oluşan veri kümelerini sınıflandırmada büyük bir başarı sağlamıştır. Konvolüsyonel ağlarda filtrelerin ve katmanların artırılması karmaşık özelliklere sahip veri kümelerinin eğitimini de mümkün hale getirmiştir. Ancak katmanların çok fazla artırılması Gradyanın Yok Olması (Vanishing Gradient) problemine neden olmuştur. Derin ağların eğitiminde hata ve hatanın türevi hesaplanarak ağırlıkları güncellemek için türev (gradyan) önceki katmanlara geri besleme yapılır. Gradyanlar en son katmandan ilk katmana doğru yansıtılırken, giderek azalır ve yok olur. Bu nedenle ilk katmanlardaki ağırlıklar hiç güncellenmemektedir veya güncelleme miktarı çok az olmaktadır. Bu durum ağın öğrenmesinin durmasına sebep olmaktadır. ResNet diğer derin ağlara benzer şekilde konvolüsyon, boyut azaltma (pooling) ve aktivasyon katmanlarını içerir. Diğer ağlardan farklı olarak gradyanı kaybetmemek için katmanların girdiyi ağırlık katmanlarının çıktısına ekleyerek ağın içinde ikisini birlikte ilerletir [20].

Şekil 2.3 ve 2.4 incelendiğinde girdiye doğru iki yönden geri besleme yapıldığı görülmektedir. Ağırlık katmanları üzerinden yapılan geri besleme derinliğin artmasıyla Gradyan Yok Olması problemine neden olur. Ancak doğrudan bağlantı

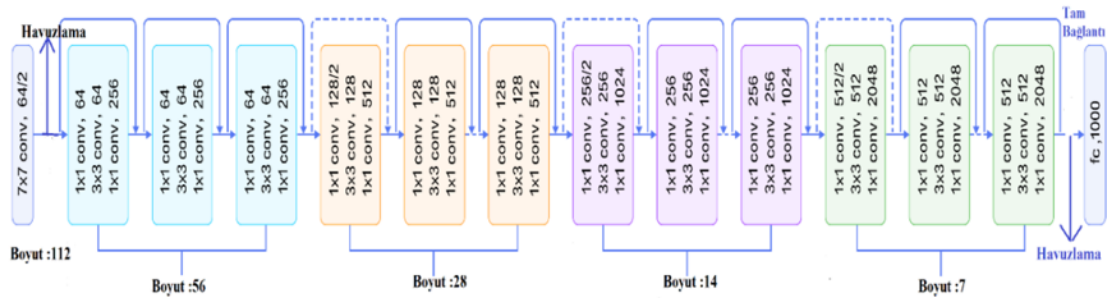


Şekil 2.3 ResNet blok yapısı [25]



Şekil 2.4 ResNet geri besleme modülü

üzerinden yapılan geri besleme ağırlık parametrelerinden bağımsız olduğu için gradyanın küçülmesini engeller. Sonuçta parametre güncellemesi ve dolayısıyla öğrenme devam eder [20].



Şekil 2.5 ResNet blok aşamaları [21]

Şekil 2.5'te gösterilen ResNet-50 mimarisi dört ana aşamadan oluşmaktadır. ResNet bloklarına geçmeden önce 7x7 filtrelerle konvolüsyon ve havuzlama (max- pooling) işlemi uygulanır. ResNet bloklarının ilk aşamasında her biri üç katmandan oluşan üç blok mevcuttur. Kıvrımlı oklar girdilerin blok çıktılarıyla doğrudan bağlantısını göstermektedir. Birinci blokta girdiye sırasıyla 64, 64 ve 256 konvolüsyonel filtre uygulanmıştır. Birinci bloğun bitiminde görülen kesik çizgili oklar girdinin boyutunun iki kat azalacağını ve özellik haritalarının ise iki katına çıkacağını göstermektedir. Bu adımlar diğer aşamalarda da tekrarlanır. Son bloğun ardından yine havuzlama yöntemi uygulanarak tam bağlantı katmanına verilir ve 1000 uzunluklu nöron vektörü elde edilir.

BigGAN çalışmasında [14]'e benzer şekilde ResNet-50 mimarisinin farklılaşmış bir biçimi kullanılmıştır. Üretici ve ayırt edici ağıın mimarisi Şekil 2.6'de özetlenmiştir.

Şekil 2.6'de ch her bir katmandaki birim sayısını veya ağıın genişliğini, BN küme normalizasyonunu, $ReLU$ aktivasyon fonksiyonunu, ResBlock up girdi boyutunu yükselten ResNet bloğunu ve ResBlock down girdi boyutunu düşüren ResNet bloğunu ifade eder.

ÜÇA sahte örnekler üretmede “doku” açısından başarısını kanıtlamıştır. Ancak, genelde yapısal özelliklerden olan “geometrik ve uzaysal” bağlamda henüz başarılı sayılmamaktadır. Örneğin üretilen bir köpek görüntüsünde köpek gerçekçi bir kürkle üretilmesine rağmen köpeğin dört bacağı belirgin ve ayık halde görünmemektedir [14]. Bunun sebebi yaygın olarak ÜÇA'nın konvolüsyonel ağlardan oluşmasıdır. Konvolüsyonel ağlar teorik olarak örnekler yapısal özellikleri çıkarmada çok başarılı bir sınıflandırıcı olarak kabul edilir. Ancak, görüntüde mekânsal olarak uzak mesafelerde bulunan birbiriyle ilişkili özelliklerin (köpeğin kafası, bacakları ve kuyruğu gibi özelliklerinin gövdeye göre yeri vb.) çıkarılabilmesi için büyük boyutlu

$z \in \mathbb{R}^{128} \sim \mathcal{N}(0, I)$	RGB image $x \in \mathbb{R}^{128 \times 128 \times 3}$
dense, $4 \times 4 \times 16 \cdot ch$	ResBlock down $1 \cdot ch$
ResBlock up $16 \cdot ch$	Non-Local Block (64×64)
ResBlock up $8 \cdot ch$	ResBlock down $2 \cdot ch$
ResBlock up $4 \cdot ch$	ResBlock down $4 \cdot ch$
ResBlock up $2 \cdot ch$	ResBlock down $8 \cdot ch$
Non-Local Block (64×64)	ResBlock down $16 \cdot ch$
ResBlock up $1 \cdot ch$	ResBlock $16 \cdot ch$
BN, ReLU, 3×3 conv 3	ReLU
Tanh	Global sum pooling
	Embed(y) $\cdot h + (\text{dense} \rightarrow 1)$

Şekil 2.6 Üretici (solda) ve ayırt edici (sağda) ağı mimari şablonları [15]

filtreler kullanılması gereklidir. Bu durum ağı büyük oranda karmaşıktır. Diğer yandan, küçük boyutlu filtreler ile ağı eğitilirken, uzak mesafelerdeki özelliklerin bağlantısını ortaya çıkarmak için çok fazla katmandan oluşan ağlar tasarlanmalı ve gizli katmanların her birindeki parametreler çok hassas şekilde ayarlanmalıdır. ÜÇA'yı diğer derin ağlardan ayıran en önemli özelliklerden biri de parametrelerin değişime aşırı duyarlı olmasından dolayı doğru ayarlamaların zorlukla yapılmasıdır. Ayrıca, parametreler eğitim kümesi için başarılı bir şekilde ayarlansa bile test kümesinde başarılı olmayabilir [14].

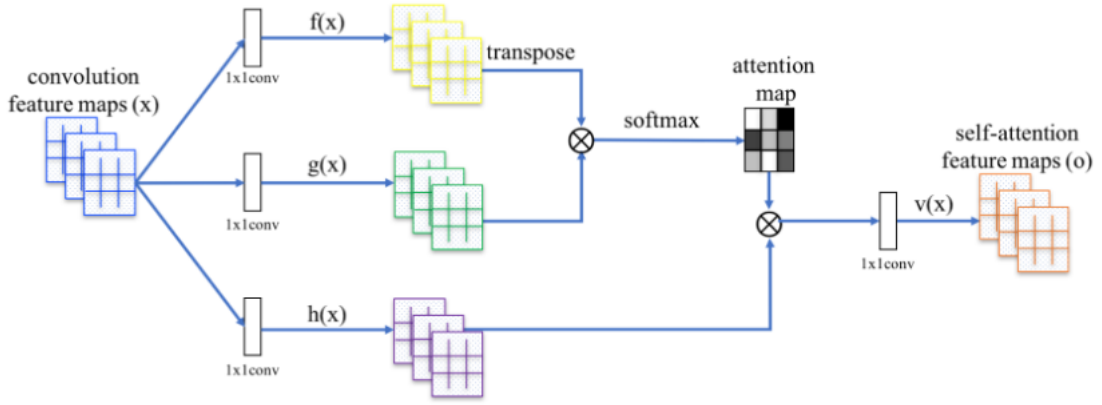
ÜÇA'da üretilen örneklerin lokal parçaları arasında bir bütünlüğün sağlanamaması problemine Öz İlgili Ağlar (Self Attention GAN, SAGAN) yeni bir yaklaşım getirmiştir. Öz-iliği prensibi bir noktadaki özelliği diğer lokasyonlardaki tüm özelliklerin ağırlıklı toplamı olarak ifade eder. Bu prensip yardımıyla ağı, ağıdaki yapısal olarak ilişkili olan lokasyonlara odaklanır.

Şekil 2.7'de önerilen yöntem özetlenmiştir. Öz-iliği modülü anahtar-değer-sorgu (key-value-query) modeline dayanır. Konvolüsyonel ağlar tarafından üretilen gizli katmandaki özellik haritasını girdi olarak alan modül, üç ayrı özellik haritası oluşturur. Bu özellik haritaları, orijinal özellik haritasının üç farklı 1×1 konvolüsyon filtresinden geçirilmesiyle elde edilmiş olan anahtar $f(x)$, değer $g(x)$, sorgu $h(x)$ 'dir. Daha sonra anahtar $f(x)$ ve değer $g(x)$ haritalarıyla matris çarpımı yapılır. Çarpım matrisinin her bir satırına softmax fonksiyonu uygulanır. Elde edilen çıktı j lokasyonundaki özelliğin

i bölgesinin üretimine ne oranda katıldığını gösterir. Sorgu özellik haritasına bunu uygulamak için ilgi haritasıyla $h(x)$ matris çarpımı yapılır ve öz ilgi haritası elde edilir.

Son adımda, (2.1)'de ilgi katmanının çıktısı olan öz ilgi haritası skaler bir değerle çarpılarak original özellik haritasına eklenir. Öz ilgi katmanı ayırt edici ve üretici ağıın her ikisine de bir kez ve gizli katman seviyesinde uygulanır.

$$y_i = \gamma o_i + x_i \quad (2.1)$$



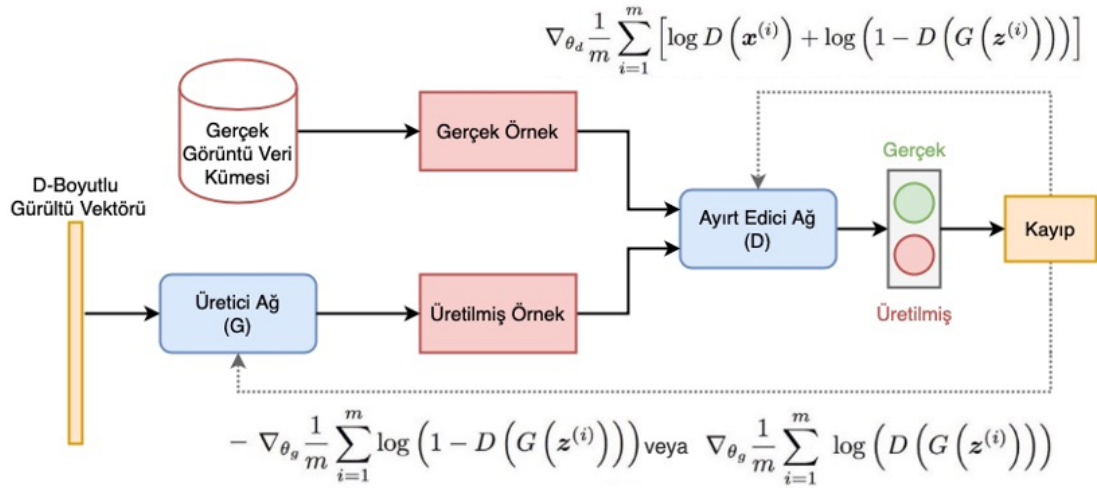
Şekil 2.7 Öz ilgi modülünün işleyişi [14]

2.3 Min-Maks Oyunu ve Kayıp Fonksiyonu

Üretici Çekişmeli Yapay Nöron Ağları (Generative Adversarial Networks), “üretici” (generator) ve “ayırt edici” (discriminator) adlı iki farklı ağdan oluşur. Üreticinin amacı rastgele oluşturulmuş gürültü girdisini örneklerle dönüştürmek; ayırt edicinin amacı, gerçek ve üretilmiş görüntü girdilerini alıp, alınan girdinin sahte veya gerçek olduğunu ayırt etmektir. Üretici ve ayırt edici, birbiriyle yarışan iki ağ olup iki oyunculu min-maks probleminde dengeyi yakalamaya çalışırlar. ÜÇA'nın temel mimarisi Şekil 2.8'de yer almaktadır.

$$\min \max \mathbb{E}_{x \sim q_{data(x)}} [\log D(x)] + \mathbb{E}_{z \sim p(z)} [\log(1 - D(G(z)))] \quad (2.2)$$

Eşitlik (2.2)'de z , $p(z)$ dağılımından üretilmiş bir örtülü değişkendir (latent variable). $p(z)$, üniform ($U[1,1]$) veya normal ($N[0,I]$) dağılım gibi bir dağılıma uyan bir rastgele değişkendir. Genellikle, G ve D konvolüsyonel ağlardır. $\mathbb{E}$ beklenen kayıp ya da toplam maliyettir. Eşitlik (2.2)'de gösterilen üretici ve ayırt edicinin hedef fonksiyonları Tablo 2.1'de belirtilmiştir [22].



Şekil 2.8 Üretici çekişmeli ağ mimarisi

Tablo 2.1 incelendiğinde $D(x)$, ayırt edici ağın x_{gercek} girdisi için $(0, 1)$ aralığında çıktı olarak verdiği olasılık değeridir. $D(G(z))$ ise ayırt edici ağın $x_{uretilmis}$ girdisi için $(0, 1)$ aralığında çıktı olarak verdiği olasılık değeridir. Ayırt edici ağ üretilmiş görüntü için 0, gerçek görüntü için 1, ayırt edemediği görüntü için 0.5 çıktısını verir. Eşitlik 2.1’de $\log(D(x))$ ayırt edici ağın x_{gercek} girdisi için kayıp fonksiyonudur. $\log(1 - (D(G(z))))$ ayırt edici ağın $x_{uretilmis}$ girdisi için kayıp fonksiyonudur.

Tablo 2.1 Üretici ve ayırt edici ağın hedef fonksiyonları

Üretici Ağda (G)	Ayırt Edici Ağda (D)
$D(G(z))$ maksimize edilmeli	$D(x)$ minimize edilmeli
	$D(G(z))$ maksimize edilmeli

$$D_{kayip} = D_{kayip_gercek} + D_{kayip_uretilmis} \quad (2.3)$$

$$D_{kayip_gercek} = \log(D(x)) \quad (2.4)$$

$$D_{kayip_uretilmis} = \log(1 - D(G(z))) \quad (2.5)$$

$$D_{kayip} = \log(D(x)) + \log(1 - D(G(z))) \quad (2.6)$$

Ayırt edici ağ gerçek veya üretilmiş görüntü girdisi için çıktı olarak $(0,1)$ aralığında

bir değer verdiği için kayıp fonksiyonu olarak $\log(x)$ tercih edilir. Eşitlik (2.4)'e göre;

$$D(x) = 1 \rightarrow \log(D(x)) = \log(1) = 0 \quad (2.7)$$

$$D(x) = 0 \rightarrow \log(D(x)) = \log(0) = \text{Tanimsiz} \quad (2.8)$$

$$\lim_{x \rightarrow 0} \log(x) = \lim_{x \rightarrow 0^+} \log(x) = -\infty \quad (2.9)$$

elde edilir. Eşitlik (2.7), (2.8) ve (2.9)'da ayırt edicinin çıktısına göre kayıp değerleri hesaplanmıştır. İdeal durumda ayırt edici gerçek görüntü girdisi için 1 çıktısını verir. Bu durumda kayıp 0'dır (bkz eşitlik (2.7)). Diğer yandan, en kötü senaryoda gerçek görüntü için 0 çıktısını verdiğinde kayıp değeri $-\infty$ değerine ıraksar (bkz eşitlik (2.8) ve (2.9)).

$$D_{\text{toplam_maliyet}} = \frac{1}{m} \sum_{i=1}^m \log(D(x_i)) + \log(1 - D(G(z_i))) \quad (2.10)$$

Eşitlik (2.10)'da ayırt edicinin m tane gerçek ve m tane üretilmiş görüntü örneği için ortalama kaybı ya da diğer adıyla toplam maliyeti hesaplanmıştır. Eşitlik(2.10)'a göre ayırt edici ağ $D_{\text{toplam_maliyet}}$ değerini maksimum yapmaya çalışırken; üretici ağ $G_{\text{toplam_maliyet}}$ değerini minimum yapmayı hedefler. $D_{\text{toplam_maliyet}}$ ve $G_{\text{toplam_maliyet}}$ fonksiyonlarında ortak değişken $\log(1 - D(G(z)))$ 'dir. Bu nedenle, üretici ağ ve ayırt edici ağ bu değişkeni minimum ve maksimum yapmada yarışır. Birbiriyle yarışan iki ağdan birisi diğerinin davranışı nasıl olursa olsun kendi davranışı değiştirmiyorsa ağlar Nash dengesine ulaşmış olur.

2.4 Kayıp Fonksiyonu

Eşitlik (2.2)'de gösterilen min-maks algoritmasından sonra ÜÇA modellerinin tasarlanmasında farklı kayıp fonksiyonları uygulanmıştır [23–25]. Wasserstein, en küçük kareler ve Hinge fonksiyonu bunlardan bazılarıdır. Bu çalışmada çekişmeli ağların karşılıklı ortalama kayıplarını hesaplamak için Destek Vektör Makineleri'nde maksimum marjini belirlemek için de kullanılan Hinge fonksiyonundan

yararlanılmıştır.

$$L_D = \mathbb{E}_{(x,y) \sim p_{data}} [\min(0, -1 + D(x, y))] - \mathbb{E}_{z \sim p_y, y \sim p_{data}} [\min(0, -1 - D(G(z), y))] \quad (2.11)$$

$$L_G = \mathbb{E}_{z \sim p_y, y \sim p_{data}} D(G(z), y) \quad (2.12)$$

Bu çalışmada birçok sınıftan örnekler üretmeyi sağlayan Koşullu Üretici Çekişmeli Ağlar'dan yararlanılmıştır. Eşitlik (2.11) ve (2.12)'de gösterilen kayıp fonksiyonlarında, eşitlik 2.1'dekinin aksine L_D ve L_G , $(z \sim p_y, y \sim p_{data})$ ve $(x, y) \sim p_{data}$ türünden fonksiyonlardır. y koşul bilgisidir ve görüntünün ait olduğu sınıfın bilgisi olarak düşünülebilir. Eşitlik (2.11) ve (2.12) incelendiğinde ayırt edici ağ L_D 'yi, üretici ağ ise L_G 'yi minimize etmeye çalışır.

2.5 Optimizasyon Problemi

Eşitlik (2.13)'de belirtildiği şekilde makine öğrenmesinde temel amaç optimizasyon problemini yani tüm eğitim örneklerinden elde edilen ortalama kaybı veya toplam eğitim maliyetini düşürmektir. $\{(x_i, y_i)\}_{i=1}^m$ ikilileri için x_i eğitim örneklerini, y_i ait olduğu sınıfları, W öğrenilecek ağırlık parametrelerini, $l(W; x_i, y_i)$ ise W parametreleriyle x_i, y_i örneği için ağın kayıp miktarını gösterir. $r(W)$ modelin karmaşıklığını azaltmak ve dolayısıyla ezberlemeyi engellemek için uygulanan regülarizasyon fonksiyonunu, $\lambda > 0$ regülarizasyon katsayısını gösterir [26].

$$\min\{L(W) = \frac{1}{m}l(W; x_i, y_i) + \lambda r(W)\} \quad (2.13)$$

Hedef fonksiyonun anlık ağırlık parametrelerine göre kısmi türevi olan gradyan, fonksiyonun ortalama kaybını arttıran yönü gösterir. Bu nedenle ortalama kaybı azaltmak için gradyanın tersi yönünde parametreler güncellenir ve bu yöntem Gradyan İnişi adı verilir. Gradyan inişi ile iyileştirmede karmaşıklık problemi ortaya çıkmaktadır. Eşitlik (2.13)'deki gibi, her iterasyonda tüm örnekler üzerinden kayıp hesaplamak (Küme Gradyan İnişi, Batch Gradient Descent) ve ağırlıkları güncellemek için ikinci derece türevi (Hessian) kullanmak bu durumun başlıca nedenleridir. Bu nedenle hedef fonksiyonunun karmaşıklığını azaltmak için birinci derece türevle güncelleme ve örnek setini küçük parçalara bölerek kayıp hesaplama (Mini-Küme Gradyan İnişi, Mini-Batch Gradient Descent) gibi çözümler uygulanmaktadır [26].

Algorithm 1 — Adam Optimizasyonu [26]

Input: $m_0 = 0, v_0 = 0$ - Birinci ve ikinci momentin ilklendirilmesi

Output: Güncellenmiş ağırlıklar

```
1: for  $t \leftarrow 1$  to  $T$  do
2:    $g_t = \nabla L_t(W_{t-1})$  ( $t$  anında  $W_j$ 'ye göre gradyan)
3:    $m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$  (Birinci moment tahmini)
4:    $v_t = \beta_2 m_{t-1} + (1 - \beta_2) g_t^2$  (İkinci moment tahmini)
5:    $\widehat{m}_t = \frac{m_t}{(1 - (\beta_1)^t)}$  (Birinci moment bias düzeltme)
6:    $\widehat{v}_t = \frac{v_t}{(1 - (\beta_2)^t)}$  (İkinci moment bias düzeltme)
7:    $W_t = W_{t-1} - \alpha \frac{\widehat{m}_t}{\sqrt{\widehat{v}_t} + \epsilon}$  (Ağırlık parametresini güncelleme)
8: end for
9: return  $W_t$ 
```

Adam optimizasyon algoritması stokastik gradyan temelli AdaGrad (Adaptive Gradients) ve RMSProp (Root Mean Square Propagation) yöntemlerini birleştiren bir algoritmadır. Bu yöntem parametre güncellemede birinci derece gradyanı kullandığı için işlem yükü azalmış olur. Adı Adaptif Moment Tahmini (Adaptive Moment Estimation)'ne dayanan Adam algoritması gradyan tahminlerinin birinci ve ikinci momentlerini kullanarak (beklenen değer ve varyans), parametreler için farklı ve adaptif öğrenme oranları belirler [26]. Her bir W_j parametresi için uygulanan adımlar Adam Optimizasyonu kısa kodunda gösterilmiştir. Kısa kodda anlaşılır olması için W_j yerine W kullanılmıştır.

$\alpha > 0$

$\beta_1 \in [0, 1)$ – Birinci moment için ağırlık azaltma oranı (genellikle 0.9)

$\beta_2 \in [0, 1)$ – İkinci moment için ağırlık azaltma oranı (genellikle 0.999)

$\epsilon > 0$ – Epsilon değeri (genellikle 10^{-8})

Yapay nöron ağının kayıp fonksiyonunun gradyanı bir rastgele değişken olarak ele alınabilir. Çünkü gradyan, veri kümesinden rastgele seçilmiş mini-küme örnekleri için hesaplanır. Kısa kodda gösterilen birinci moment gradyan değişkeninin beklenen değerini, ikinci moment ise varyansını ifade eder. Adam optimizasyon algoritması momentleri tahmin ederken o anki mini-küme örnekleri için gradyanın ve gradyanın karesinin hareketli ortalamalarını kullanır. Kısa kod (1) incelendiğinde, m_t ve v_t ilk adımda sıfır değeriyle ilklendirilir. Bu nedenle özellikle ilk iterasyonlarda ve azaltma oranları 1'e yakın olduğu için sıfıra karşı bias eğilimindedirler. Bu nedenle parametrelerin güncellenmesinden önce momentlere bias düzeltme uygulanır [26].

α öğrenme oranı ağırlık parametreleri güncellenirken ağırlıklarda yapılacak değişimin küçük ya da büyük olmasını belirler. Modelin öğrenmesini hızlandırır ya da yavaşlatır. Eğitimde her bir adımda az sayıda örnek kullanılıyorsa varyans yükselir

ve ezberlemeye sebep olur. Bu durumu engelleyerek modelin genelleme yeteneğini arttırmak için bir regülarizasyon çeşidi olan ağırlık azaltma oranı (weight decay) uygulanmıştır.

2.6 Regülarizasyon Teknikleri

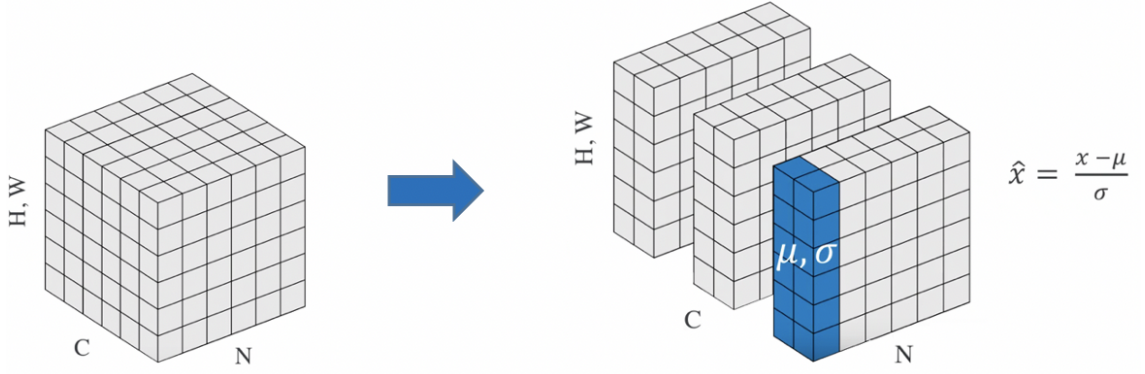
Makine öğrenmesinde geliştirilen algoritmaların odaklandığı temel sorun genelleştirme problemidir. Bu problemin kaynağı algoritmaların öğrenememesi veya ezberleme yapmasıdır. Bu nedenle algoritmalar eğitim hatası yerine genelleştirme hatasını azaltmak için modellenir. Genelleştirilmiş algoritmalar hem eğitim hem de test veri kümesinde benzer şekilde başarılı sonuçlar verir [27]. Bu nedenle derin ağların eğitiminde de genelleştirme hatasını azaltmak için uygulanan birçok yöntem mevcuttur. Regülarizasyon teknikleri adı verilen bu yöntemlerden bazıları, ağa kısıtlayıcı parametreler veya ceza parametreleri verme ve veri kümesini genişletmedir. Bu çalışmada kullanılan regülarizasyon teknikleri aşağıda açıklanmıştır.

2.6.1 Normalizasyon

Nöron ağırları eğilirken, ilk katmana verilen girdileri standart hale getirmek için girdilere normalizasyon işlemi uygulanır. Benzer şekilde iç katmanlardaki girdilerde de normalizasyona ihtiyaç duyulmuştur. Çünkü bir katmandaki nöronların parametreleri değiştikçe aktivasyonları, dolayısıyla da onlara bağlı olan sonraki katmanların girdilerinin dağılımı değişir. Bu durum eğitim esnasında kararsızlığa sebep olmakta ve eğitimi zorlaştırmaktadır. [15] çalışmasında, nöron aktivasyonlarına uygulanan Küme Normalizasyonu (KN) (Batch-Normalization) yöntemiyle bu probleme çözüm getirilmiştir. KN'nin eğitimi hızlandırması ve başarıyı artırması için küme büyüklüğü yüksek olmalıdır [28]. Bu çalışmada kullanılan kaynaklar ile küme büyüklüğü yeterli miktarda arttırılamadığı için KN eğitim başarısını kötüleştirmiştir. Bu nedenle normalizasyon katmanında Grup Normalizasyonu (GN) (Group-Normalization) kullanılmıştır. KN, küme büyüklüğünün 32'den düşük olduğu durumlarda başarısızdır. GN ise küme büyüklüğü 2'ye kadar düşürüldüğünde dahi verimli çalışmaktadır. Yapılan çalışmada küme büyüklüğü 26 olarak alınmıştır. GN Şekil 2.9' da gösterilmiştir.

$$\widehat{x}_i = \frac{1}{\sigma_i}(x_i - \mu_i) \quad (2.14)$$

Eşitlik (2.14)'te x normalize edilecek özelliği ve i özelliğin indisini temsil eder. GN katmanının girdisi ve çıktısı olan $i = (i_N, i_C, i_H, i_W)$ 4 boyutlu vektörlerdir.



Şekil 2.9 Grup normalizasyonu [28]

N, C, H, W sırasıyla, i . özelliğin küme büyüklüğü, kanal, uzaysal yükseklik ve genişlik bilgisidir. Girdinin görüntü olduğu varsayıldığında kanallar konvolüsyonel ağdaki özellik haritalarıdır.

GN katmanı kanalları gruplara böler ve her bir grubun içindeki aktivasyonları normalize eder. Eşitlik (2.14)'te GN, grubun ortalama aktivasyonunu μ girdi aktivasyonlardan çıkarak, grubun aktivasyonlarının standart sapmasına σ bölerek normalizasyonu gerçekleştirir. Gruptaki aktivasyonların ortalaması ve standart sapması eşitlik (2.15)'teki gibi hesaplanır.

$$\mu_i = \frac{1}{m} \sum_{k \in S_i} x_k, \quad \sigma_i = \sqrt{\frac{1}{m} \sum_{k \in S_i} (x_k - \mu_i)^2 + \epsilon} \quad (2.15)$$

Eşitlik (2.15)'te S_i piksel setini, ϵ önemsiz sabiti ve m setin büyüklüğünü ifade eder. Eşitlik (2.16)'te piksel seti hakkında bilgi verilmektedir.

$$S_i = \{k | k_N = i_N, \lfloor \frac{k_C}{C/G} \rfloor = \lfloor \frac{i_C}{C/G} \rfloor\} \quad (2.16)$$

Eşitlik (2.16)'te grup sayısını, C/G her bir gruptaki kanal sayısıdır, $\frac{k_C}{C/G}$ 'de i ve k indisleri aynı grupta yer alan kanallardır ve her grup kanal ekseninde boyuna art arda sıralı olarak tutulur. GN μ ve σ değerini her grup için C/G kanal için hesaplar. Kısaca, aynı grupta yer alan piksellerin ortalaması ve standart sapması aynıdır. Şekil 2.9'de her biri 2 kanaldan $C/G = 2$ oluşan 3 grup $G = 3$ için basit bir GN hesaplama örneği gösterilmiştir.

2.6.2 İki Zamanlı Güncelleme Kuralı

Temelde Eşitlik (2.2)'deki prensip üzerine kurulan ÜÇA ilk ortaya konduğu günden bugüne çeşitli modifikasyonlarla evrilmiştir. ÜÇA stokastik gradyan temelli optimizasyon teknikleriyle eğitilir. Ancak optimizasyon problemlerinin yanında, diğer derin ağlardan farklı olarak ÜÇA eğitiminin başarısı, üretici ve ayırt edicide kullanılan kayıp fonksiyonuna, model mimarisine, regülarizasyon ve normalizasyon yöntemlerinin her birine karşı çok hassastır [29]. Bu çalışmada kayıp fonksiyonu olarak [14] çalışmasına benzer şekilde Hinge Kayıp Fonksiyonu kullanılmıştır. Ağırlıklar Ortogonal Regülarizasyon ile ilklendirilmiştir. Eğitim esnasında İki Zamanlı Güncelleme Kuralı (Two- Timescale Update Rule, TTUR) [30] uygulanmıştır. Bu kural gereği üretici ve ayırt edici ağa farklı öğrenme oranları verilmiştir. Farklı kombinasyonlar denendikten sonra üretici ağ için 0.0001 ve ayırt edici ağ için 0.0004 oranlarıyla modelin verimli çalıştığı görülmüştür. Böylece üretici ayırt ediciyi kandırmak için daha küçük ve temkinli adımlarla ilerlemektedir.

2.7 Performans Değerlendirme Metrikleri

Bu çalışmada ÜÇA'nın performansını değerlendirmek için Başlangıç Skoru (BS) (Inception Score, IS) [31] ve Fréchet Başlangıç Mesafesi (FBM) (Fréchet Inception Distance, FID) [32] metriklerinden yararlanılmıştır. Başlangıç Skoru, ImageNet üzerinde önceden eğilmiş Inception Net yardımıyla üretilmiş görüntülerin görsel kalitesini ve çeşitliliğini ölçer. Üretilen örneklerin olasılıksal dağılımları arasındaki Kullback-Leibler Uzaklığı (KL Uzaklığı) (Kullback-Leibler Divergence)'nı hesaplar. Aynı sınıftaki örneklerin dağılımının düşük entropiye sahip olması gerekir. Düşük entropi örneklerin kaliteli olduğunu gösterir. Tüm sınıflardaki örneklerin marjinal dağılımının yüksek entropiye sahip olması ise örneklerin birbirinden farklı olduğunun göstergesidir. Ancak, Başlangıç Skoru sadece üretilen örneklerin dağılımlarını ele aldığı için çeşitliliği doğru değerlendirse de görsel kaliteyi ölçmede yeterli değildir [30]. Çünkü ezberlemeyi (overfitting) tespit edemez ve görüntü çözünürlüğüne karşı çok hassastır. Bunun için örneklerin kalitesini değerlendirmek için BS'nin yanında Fréchet Başlangıç Mesafesi'nden yararlanılmıştır.

FBM üretilmiş ve gerçek örneklerin her ikisini de Inception Ağı'na vererek, çıktı katmanından hemen önce elde edilen aktivasyonları Gauss Dağılımı'na benzetir. Sonra, iki Gauss eğrisi arasındaki mesafeyi hesaplar. Böylece üretilen örneklerin gerçek örneklerden ne kadar uzak olduğunu ölçer. Özetle, FBM değerinin düşük; BS değerinin yüksek olması modelin başarılı olduğunu gösterir.

3

DENEYSEL SONUÇLAR

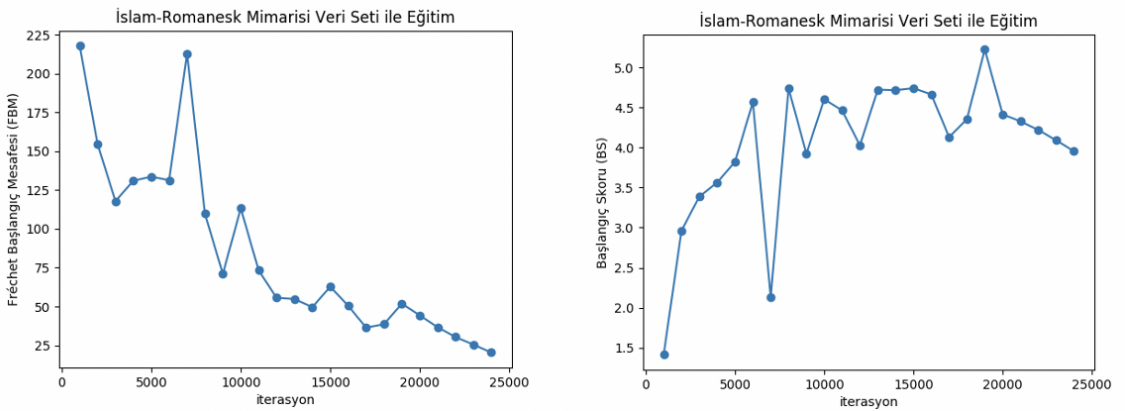
Derin öğrenme tabanlı üretici modellerden birisi olan Üretici Çekişmeli Ağlar görüntü örnekleri üretiminde 2014 yılından bugüne büyük bir ilerleme kaydetmiştir. ÜÇA dağ, gökyüzü ve deniz gibi detay içermeyen, az sayıda parçalardan oluşan ve yumuşak geçişli görüntü örnekleri üretmede başarısını kanıtlamıştır. ImageNet gibi çok sınıflı, daha karmaşık ve görece yüksek çözünürlükteki örnekler içeren veri kümelerinde ise son yıllara kadar yeterince iyi sonuçlar elde edilememiştir. Çünkü, küçük boyutlardaki konvolüsyonel filtrelerden oluşan ÜÇA, görüntülerdeki nesneleri oluşturan parçalar arasında semantik bir bağlantı yakalayamamaktadır. Ağ birbirine yakın olan piksel bölgelerinde özellikleri çıkarabilmesine rağmen birbirine uzaysal olarak uzakta bulunan ve birbiriyle semantik ilişkisi olan bölgeleri bağımsız değerlendirmektedir. Bu problem [14] çalışmasında ele alınmıştır. Bir bölgedeki özelliği etkileyen, diğer bölgelerdeki özellikleri tespit eden Öz İlgili Ağlar parça-bütün ilişkisine semantik bir yaklaşım getirmiştir. BigGAN [15] çalışması da bu semantik yaklaşımı daha büyük modeller oluşturarak ve regülerizasyon teknikleriyle geliştirmiştir. Sonuçta, [15] ile yüksek çözünürlüklü, yapısal ve dokusal olarak başarılı örnekler üretme konusunda kilometre taşı sayılabilecek aşamalar kaydedilmiştir.

Bu çalışmada BigGan'ın farklı mimari tarzlarda bina üretimi konusundaki başarısı araştırılmıştır. Yaklaşık olarak, 114000 Romanesk ve 120000 İslam mimarisi görüntüsü ile eğitilen ağın genel dış cephe, iç mekan, kubbe, çan kulesi, mukarnas ve vitray gibi mimari unsurları üretmedeki başarısı incelenmiştir. Öz İlgili Ağlar tabanlı modelin, yapısal, dokusal ve detaylı özellikler arasındaki parça-bütün ilişkisini yakalama başarısı değerlendirilmiştir. Ayrıca, modelin eğitim başarısını arttırmak için model üzerinde yapılan düzenlemeler ve eğitime etkileri performans değerlendirme metrikleriyle açıklanmıştır. 48 kişinin katılımıyla bir anket düzenlenerek üretilen örneklerin gerçekçiliği değerlendirilmiştir. Ayrıca mimaride uzman 24 kişinin katılımıyla düzenlenen anket Romanesk mimarisi örneklerinin İslam mimarisi örneklerine göre daha başarılı olduğunu göstermiştir. Model gerçekçi görüntüler ürettiği için ağın ezberleme yapmış olması olasılığı da bu bölümde değerlendirilmiştir.

3.1 Değerlendirme Metrikleri ile Deneysel Sonuçların İncelenmesi

Bu kısımda, model üzerinde yapılan düzenlemelerin etkileri Fréchet Başlangıç Mesafesi ve Başlangıç Skoru performans metrikleri açısından değerlendirilmiştir. Önceki çalışmalarda [15, 33] mini-küme (mini-batch) büyüklüğünü arttırmak öğrenmeyi hızlandırmakta ve BS değerini arttırmaktadır. Çünkü mini-küme büyüklüğü arttıkça daha çok örnek gradyan hesaplamalarının gürültüye dayanıklı olmasını ve modelin dağılımı genelleştirmesini sağlıyor. Ancak bellek kaynakları yeterli olmadığı için büyüklüğü 128x128 boyutlu örnekler için en fazla 26 olarak alınmıştır. BS'nin düşük olma sebebi mini-küme büyüklüğüdür. 256x256 boyutlu örneklerde ise küme büyüklüğü 2'den daha yükseğe çıkarılamamıştır. Sonuçta FBM 200 değerinin altına düşmemiştir ve BS 3 değerini aşamamıştır. Bu nedenle model eğitimi ve parametre ayarlamaları 128x128 boyutlarındaki örneklerle yapılmıştır.

Model iki sınıftan oluşan 128x128 boyutlarına dönüştürülmüş, yaklaşık olarak 39000 örnekli veri kümesiyle eğitilmiş ve 168 FBM değeri elde edilmiştir. Orijinal görüntüler yüksek çözünürlüklerde olduğu için yeniden boyutlandırmanın yol açtığı bilgi kaybını azaltmak amacıyla, her bir görüntüden orijinali ve 5 farklı rastgele görüntü kırılmış ve yeniden boyutlandırılmıştır. Yaklaşık olarak 234000 görüntüyü içeren artırılmış veri kümesi ile model yeniden eğitildiğinde FBM değeri 67'ye düşmüştür. Derin ağlarda derinliği arttırmak kompleks veri kümelerinde eğitim başarısını arttırmaktadır. Bu nedenle modelin derinliği iki katına çıkarılmış ve FID değeri 45'e kadar düşmeyi başarmıştır. Mini-küme büyüklüğü 26 olduğu için, Küme Normalizasyonu yöntemi modeli eğitmede başarılı olamamıştır. Bunun yerine Grup Normalizasyonu kullanılmıştır. GN eğitimin başarısına büyük bir katkı sağlamış ve FBM değeri 20'ye düşürülmüştür. Mod çöküşünden hemen önce kaydedilen FBM ve BS değerleri Şekil 3.1'de incelenebilir.



Şekil 3.1 Eğitilen modelin Fréchet Başlangıç Mesafesi ve Başlangıç Skoru sonuçları (sırasıyla)

3.2 Üretilen Yapı Örneklerinin Mimari Açıdan Değerlendirilmesi

ÜÇA ile üretilen Romanesk ve İslam mimarisi örnekleri yapısal, dokusal ve detay özellikler açısından farklı yönleriyle değerlendirilmiştir. Mimari yapıların dış cephesi bir bütün olarak ele alındığında yapının tarzını gösteren geometrik unsurlar yapısal özellik olarak ele alınmıştır.

3.2.1 Romanesk Mimarisi Örneklerinin Mimari Açıdan Değerlendirilmesi

Romanesk mimarisi kalın duvarlarla oluşturulmuş dış cephe, yüksek tonoz ve kemerlerden oluşan geniş iç mekanları içerir. Yarım daire şeklinde apsislerle oluşturulmuş parçalı dış cephe ve apsislerin yuvarlaklığını tamamlayan konik çatı, uzun ve köşeli çan kuleleri mimarinin tarzını yansıtan temel unsurlardır. Diğer yandan, dikdörtgen biçimli bir plana oturtulmuş köşeli yapılar da bu mimari tarzda yaygındır. Diğer mimari tarzlarla karşılaştırıldığında dış ve iç mekân detay içermeyen, düz, sade ve sağlamlığın ön planda olduğu bir tasarıma sahiptir. Kapı, pencere ve kemerler silindriktir. Binalarda genellikle küçük pencereler kullanılmıştır [34].

Romanesk mimarisinde taş malzemeden oluşan dış ve iç cephe mevcuttur. Dış cephede kapı ve pencereler karmaşık olmayan doku süslemeleri, yapay kemer ve sütunlarla bezenmiştir. Romanesk mimarisi mütevazı bir dokuya sahiptir. Romanesk mimarisi sade bir tarza sahip olduğu için detay unsurlar yalnızca vitray süslemeleri ve kapı girişlerindeki heykel insan figürleri olarak değerlendirilmiştir.

3.2.1.1 Dış Cephe Örneklerinin Değerlendirilmesi

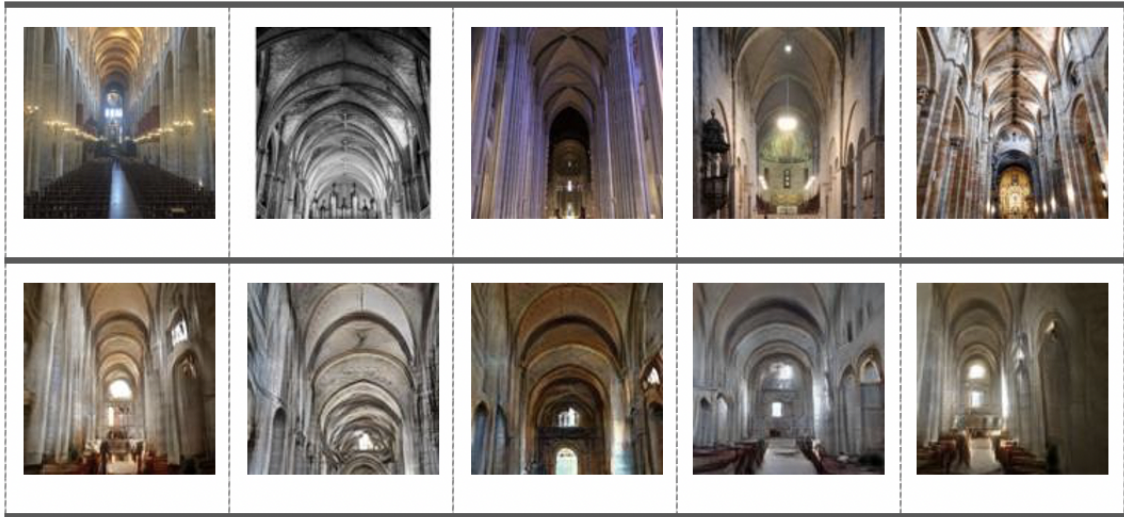


Şekil 3.2 Romanesk mimarisine ait gerçek (üstte) ve üretilen (altta) dış cephe örnekleri

Şekil 3.2’de gerçek ve üretilen dış cephe görüntü örnekleri görülmektedir. Yapısal olarak incelendiğinde, dikdörtgen veya yarım daire apsislerden oluşan dış cephe, uzun çan kuleleri, yuvarlak pencere, kapı ve kemerler model tarafından başarılı bir şekilde oluşturulmuştur. Ayrıca genel dış cephe, çan kulesi, pencere, kapı girişi ve kemerler parça-bütün ilişkisi açısından değerlendirildiğinde, mekân uzayında doğru yerlerde konumlanmıştır. Öz İlgi Ağları birbirinden uzak bölgelerde yer alan ve semantik olarak bağımlı unsurları tespit ederek, parçaları üretme konusunda başarılı olduğunu göstermiştir.

3.2.1.2 İç Mekan Örneklerinin Değerlendirilmesi

Şekil 3.3’te birinci satırda gerçek ve ikinci satırda üretilen iç mekân tonoz ve sunak görüntü örnekleri görülmektedir. Üretilen görüntülerde yüksek sütunların ve tonozların oluşturduğu geniş ve sade iç mekân gerçekçi bir şekilde tasarlanmıştır.

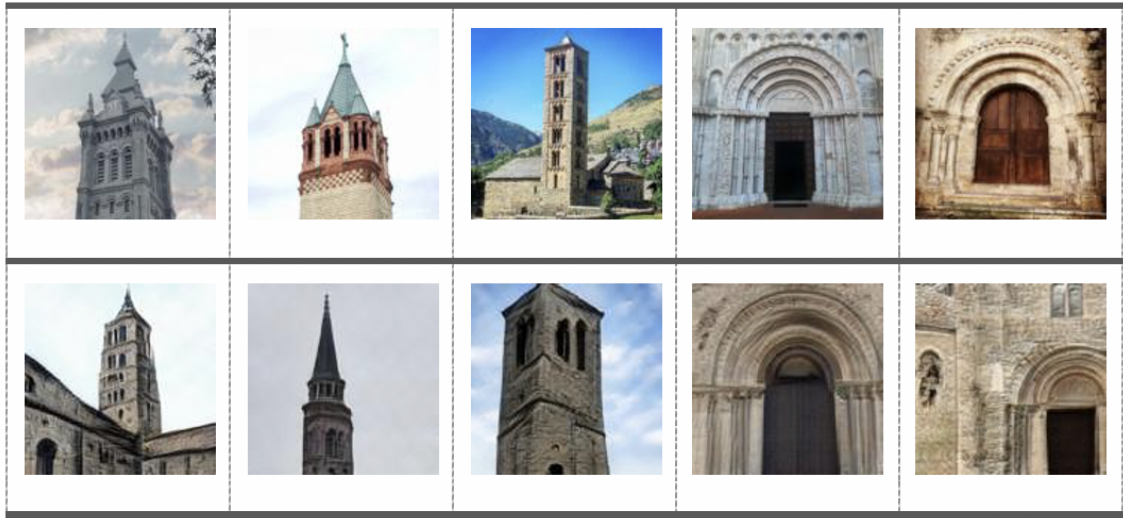


Şekil 3.3 Romanesk mimarisine ait gerçek (üstte) ve üretilen (altta) tonoz ve sunak örnekleri

3.2.1.3 Mimari Yapı Unsurlarının Değerlendirilmesi

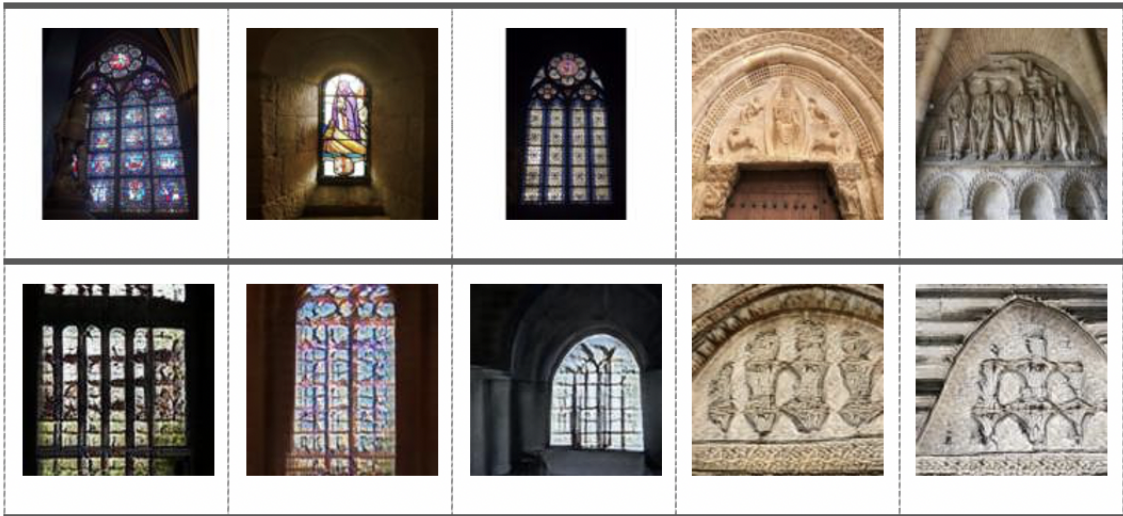
Şekil 3.4’te birinci satırda gerçek ve ikinci satırda üretilen çan kulesi ve kapı girişi görüntü örnekleri görülmektedir. Yapay kemer ve sütunlarla süslenmiş kapı girişlerinde dokusal ve yapısal özellikler başarılı bir şekilde üretilmiştir. Çan kuleleri bir bütün olarak Romanesk tarzını yansıtmaktadır. Pencere ve kapı girişleri gerçek örneklerle benzer şekilde silindirik ve küçük tasarlanmıştır.

Şekil 3.5’te birinci satırda gerçek ve ikinci satırda üretilen, vitray ve kapı girişi heykel figürlerinin örnekleri mevcuttur. Model çok detay içeren mimari unsurları üretmede yeterli değildir. Vitray ve kapı girişini şekil ve doku açısından doğru



Şekil 3.4 Romanesk mimarisine ait gerçek (üstte) ve üretilen (altta) çan kulesi ve kapı girişi örnekleri

oluşturmasına rağmen detayları yakalayamamıştır. Eğitilen görüntü örneklerinin 128x128 boyutunda olması bu durumun sebeplerinden birisi olabilir. Ayrıca süslemelerin insan gözüyle ayırt edilmesi dahi çok güçtür.



Şekil 3.5 Romanesk mimarisine ait gerçek (üstte) ve üretilen (altta) vitray ve kapı süslemesi örnekleri

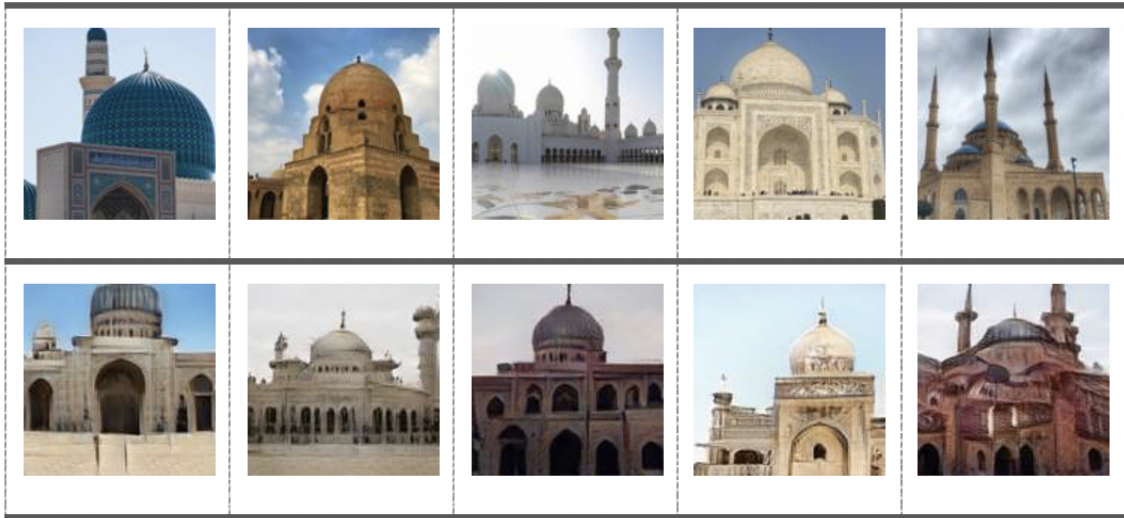
3.2.2 İslam Mimarisi Örneklerinin Mimari Açıdan Değerlendirilmesi

İslam mimarisi tarzına sahip yapılar geniş bir coğrafyada konumlandığı için farklı kültürel özelliklerin bir araya gelmesiyle oluşmuştur. Bu sebeple İslam mimarisi Romanesk mimarisinden daha çeşitli ve karmaşık bir tarza sahiptir. Yapısal özellikleri açısından İslam mimarisi kare veya dikdörtgen biçiminde bir avlu içinde yer alan kalın duvarlarla oluşturulmuş ana strüktür ve genellikle birden çok daire veya yarım daire

şeklinde kubbelerden oluşur. Ana strüktürün genellikle köşelerine konumlandırılmış olan minareler mevcuttur. Avlu revaklarının ön yüzünde sıralı kemerler bulunur. Üstü tonoz veya kubbelerle örtülü revak yapının duvarına dayanır. Mimaride yaygın unsurlardan birisi de uzun sütunlar ve sivri bir kemerle oluşturulmuş, üç tarafı kapalı ve bir tarafı açık bir yapı parçası olan eyvandır. İslam mimarisi unsurları bir bütün olarak simetrik bir yapı oluşturur [35].

İslam mimarisinde yapı malzemesi taş, kerpiç, mermer veya tuğla olarak farklılık gösterir. Bu durum mimari yapıların dokusunu farklı göstermektedir. Çini, altın ve pirinç görünümlü duvarlar, kubbeler dış mekânda yaygın olarak kullanılan unsurlardandır. Kapı girişinin ve eyvanın üst kısmında mukarnas adı verilen çokgenlerle oluşturulan girintili çıkıntılı geometrik bezeme unsurları yer alır. İç ve dış mekân süslemelerinde geometrik desenler ve çiçek desenleri kullanılır. Bu süslemeler özellikle çinilerin üstünde bezenmiştir.

3.2.2.1 Dış Cephe Örneklerinin Değerlendirilmesi



Şekil 3.6 İslam mimarisine ait gerçek (üstte) ve üretilen (altta) dış cephe örnekleri

Şekil 3.6’da birinci satırda gerçek ve ikinci satırda üretilen, İslam mimarisi dış cephe görüntü örnekleri görülmektedir. Yapısal olarak üretilen örneklerde minare, kubbe, eyvan, kapı ve pencere unsurları doğru biçimde üretilmiştir. Bu unsurlar parça-bütün ilişkisi açısından değerlendirildiğinde üretilen örnekler tutarlıdır.

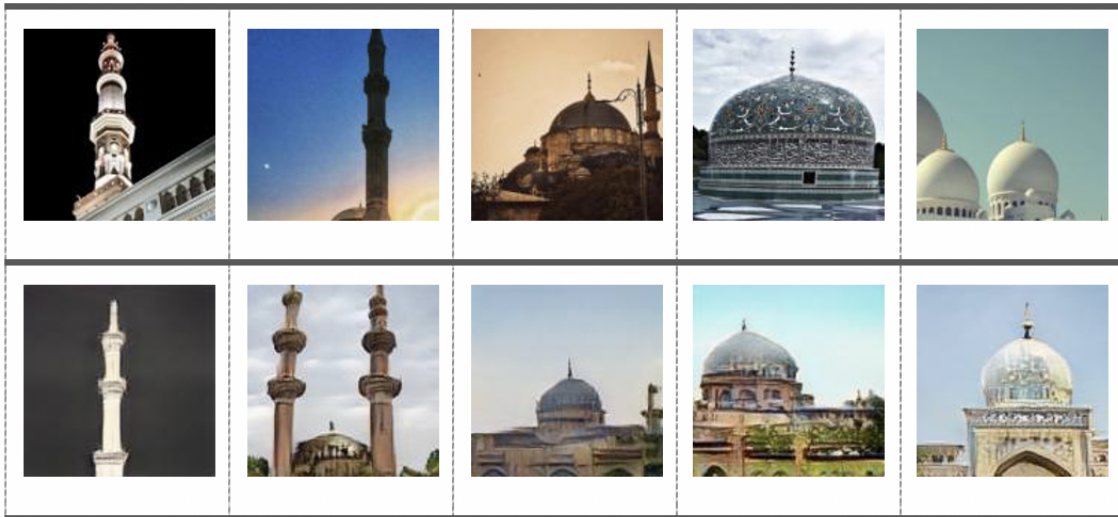
3.2.2.2 Mimari Yapı Unsurlarının Değerlendirilmesi

Şekil 3.7’de birinci satırda gerçek ve ikinci satırda üretilen, eyvan görüntü örnekleri görülmektedir. Yapısal olarak sivri kemerlerle oluşturulmuş olan eyvan ve kapı girişleri

gerçekçi örneklerdir. Ancak doku özellikleri ele alındığında, model yetersiz kalmıştır. Bu durum eğitilen görüntü örneklerinin boyutundan kaynaklanıyor olabilir.



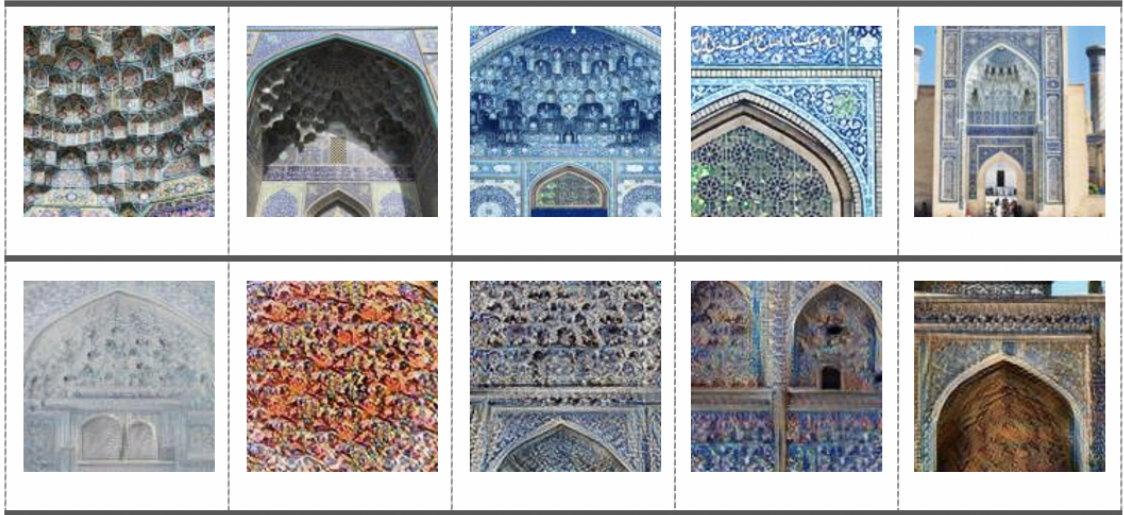
Şekil 3.7 İslam mimarisine ait gerçek (üstte) ve üretilen (altta) eyvan örnekleri



Şekil 3.8 İslam mimarisine ait gerçek (üstte) ve üretilen (altta) minare ve kubbe örnekleri

Şekil 3.8’de birinci satırda gerçek ve ikinci satırda üretilen, minare ve kubbe görüntü örnekleri görülmektedir. Genellikle doku ve ayrıntı içermediği için model, temel İslam mimarisi unsurları olan kubbe ve minare üretme konusunda başarılıdır.

Şekil 3.9’da birinci satırda üretilen ve ikinci satırda gerçek, mukarnas ve doku özelliği içeren görüntü örnekleri görülmektedir. Model detaylı özellikleri üretmede yeterli değildir. Mukarnasın bulunduğu kapı girişlerini geometrik olarak oluşturabilmesine rağmen, girintili ve çıkıntılı detayları yakalayamamıştır. Ayrıca, sağdaki iki görüntüde gösterilen eyvan örneklerinde çini dokusunu renklerle oluşturulmuştur. Ancak model çini süsleme detaylarını yakalama konusunda başarılı değildir.



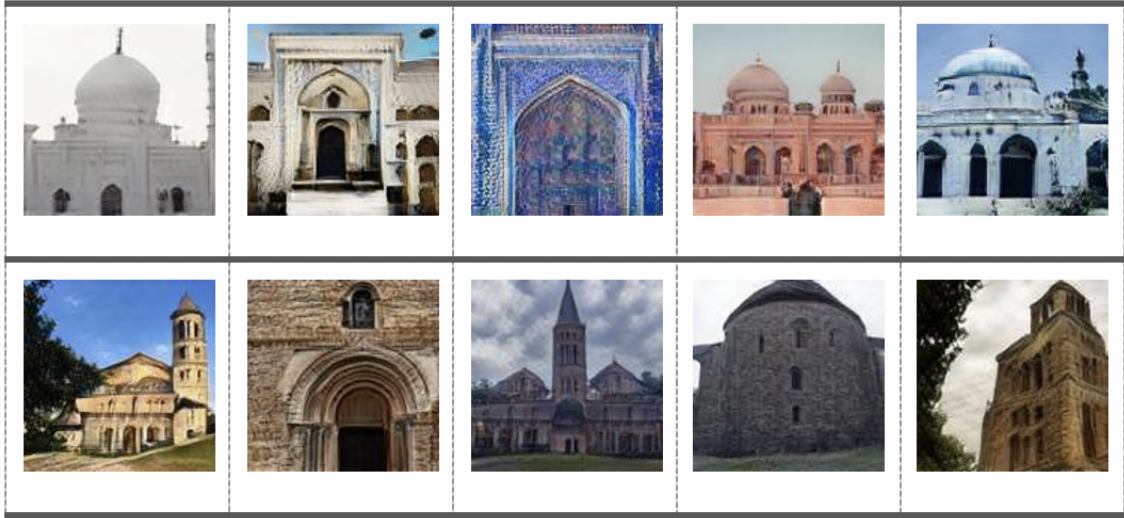
Şekil 3.9 İslam mimarisine ait gerçek (üstte) ve üretilen (altta) mukarnas ve süsleme örnekleri

Özetle model Romanesk ve İslam mimarisi örneklerini üretmede genel olarak başarılıdır. Yapısal ve dokusal özellikler ağ tarafından doğru biçimde oluşturulabilmektedir. Öz İlgili Ağları, mimari yapıların temel özelliklerinden birisi olan bütün ve parçaların birbirini tamamlayıcılığı konusunda başarılı olduğunu göstermiştir. Ancak model mimari unsurların detaylarını yakalamada yetersizdir. Bu durum eğitilen görüntülerin düşük boyutta (128x128) olmasından kaynaklı olabilir. Yeniden boyutlandırma yapmadan görüntülerin orijinal boyutlarıyla ele alınması mimari detay özelliklerini çıkarmada faydalı olur. Ancak, yüksek boyutta görüntü örnekleri ağı karmaşıklığını arttırmaktadır. Bu nedenle, orijinal görüntüler parçalara bölünerek eğitilirse daha gerçekçi ve kaliteli örnekler üretilir. Ekler kısmında üretilen örnekler incelenebilir. a

3.3 Üretilen Yapı Örneklerinin Uzman Tarafından Değerlendirilmesi

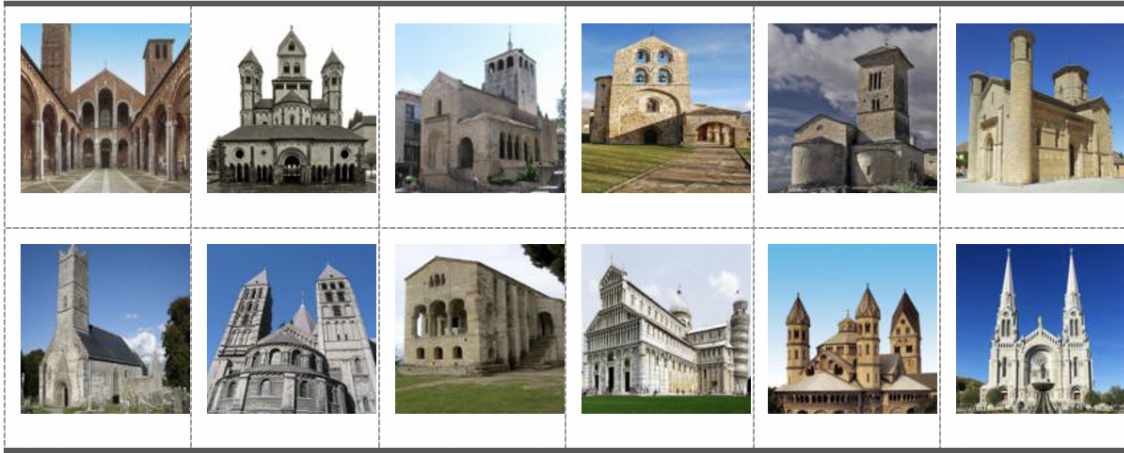
Üretilen görüntü örnekleri bir mimari uzmanı tarafından teknik açıdan değerlendirilmiştir. Bu bölümde modelin başarısını değerlendirmek amacıyla uzman görüşlerine yer verilmiştir.

Üretilen örneklerde İslam mimarisi örneklerinin görüntü kalitesi Romanesk mimarisi örneklerine göre daha düşüktür (Şekil 3.10). İslam mimarisi Asya, Avrupa ve Afrika kıtaları içeren çok geniş bir coğrafyaya ve yedinci yüzyıldan günümüze uzanan büyük bir zaman dilimine yayıldığı için malzeme, üslup, biçim ve doku farklılığı yapı örneklerinin çok çeşitlenmesine neden olmaktadır (Şekil 3.11). Bu nedenle



Şekil 3.10 Üretilen İslam mimarisi (üstte) ve Romanesk mimarisi (altta) görüntü örnekleri

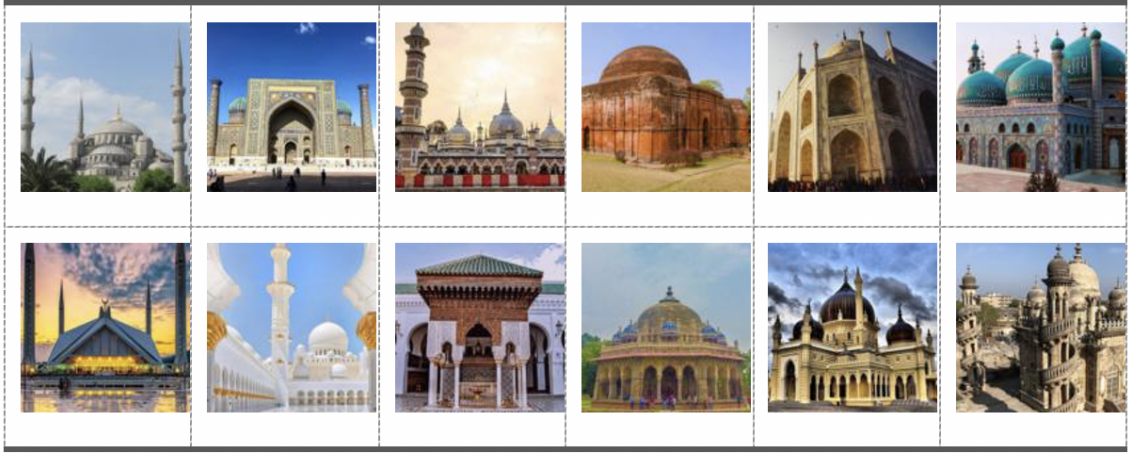
İslam mimarisi kategorisi modeli karmaşık hale getirmektedir. Bu durum modelin öğrenmesini zorlaştırarak üretilen İslam mimarisi örneklerinin daha az gerçekçi olmasına yol açmaktadır. Doğu Avrupa’da dört yüzyıl hakim olan Romanesk mimarisi ise coğrafya ve tarihsel dönem olarak dar bir alanı kapsadığından dolayı, mimari üslup olarak da diğer stillerden kesin çizgilerle ayrıldığı için eğitim örnekleri birbirine benzerdir (Şekil 3.12). Bu durum modelin Romanesk mimarisinde örnek üretme başarısını arttırmıştır.



Şekil 3.11 Romanesk mimarisi gerçek görüntü örnekleri

3.3.1 Üretilen Örneklerde Başarısız Özelliklerin Değerlendirilmesi

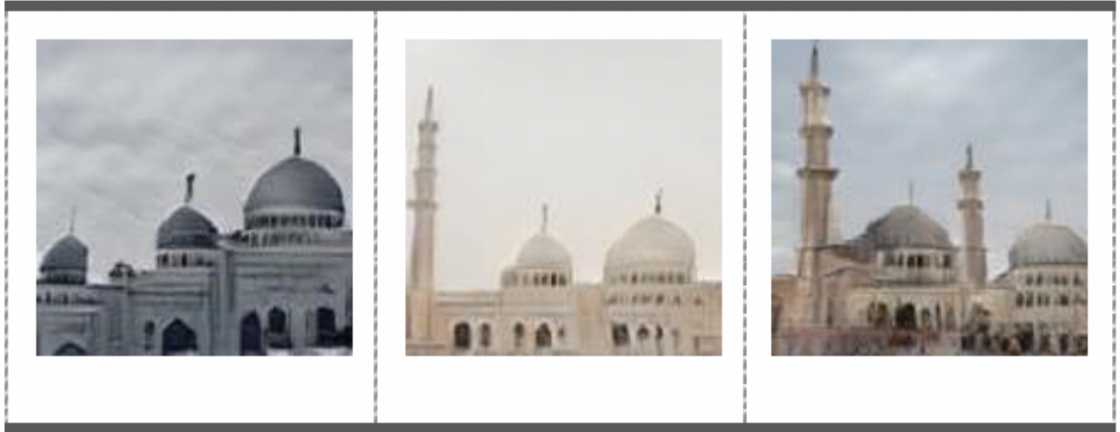
Romanesk ve İslam mimarisinde örneklerin çoğunluğunda saptanan genel hatalar bu bölümde açıklanmıştır.



Şekil 3.12 İslam mimarisi gerçek görüntü örnekleri

3.3.1.1 İslam Mimarisi Yapı Unsurlarının Değerlendirilmesi

Eşit büyüklükte büyük kubbelerden oluşan yapılar merkezi kubbeli ve çevresel küçük kubbelerden oluşan islam mimarisi üslubuna uyumlu değildir (Şekil 3.13).



Şekil 3.13 Üretilen hatalı kubbe örnekleri

Eğitimde kullanılan örnekler farklı üsluplarda olduğu için soğan, konik ve yarım daire gibi çok çeşitli kubbe örneklerini içermektedir. Bu nedenle üretilen örneklerde Şekil 3.14'teki gibi kubbe formu düzgün biçimde oluşturulmamıştır.

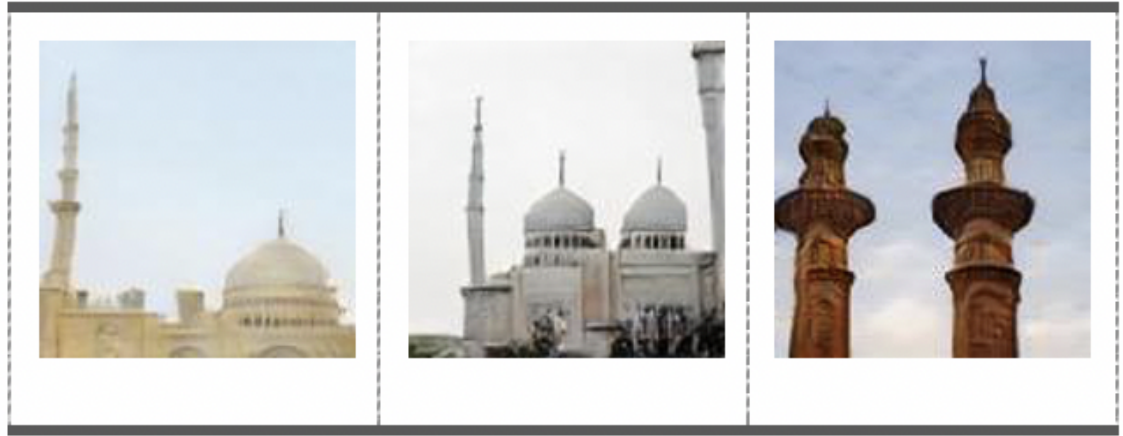
Şekil 3.15'de görüldüğü gibi minare formu bir doğrultuda değildir ve yaslama mevcuttur.

3.3.1.2 İslam Mimarisi Dış Cephe Örneklerinin Değerlendirilmesi

Şekil 3.16 incelendiğinde uzaktaki ve yakındaki nesnelerin büyüklüğünü belirleyen perspektif bilgisi doğru yansıtılmamıştır.



Şekil 3.14 Üretilen hatalı kubbe örnekleri



Şekil 3.15 Üretilen hatalı minare örnekleri



Şekil 3.16 Üretilen hatalı İslam mimarisi örnekleri

3.3.1.3 Romanesk Mimarisi Dış Cephe Örneklerinin Değerlendirilmesi

Romanesk mimarisine ait örnekler İslam mimarisi örneklerine göre daha başarılı olsa da Romanesk yapıların da göze çarpan eksiklikleri bulunmaktadır.



Şekil 3.17 Üretilen Romanesk mimarisi örnekleri

Şekil 3.17’de görüldüğü üzere, Romanesk mimaride çan kulesi ana girişin sağında veya solunda olmalıdır. Ancak üretilen örneklerde çan kulesi ana girişin hemen arkasında yer almaktadır.



Şekil 3.18 Romanesk mimari üretilen dış cephe örnekleri

Şekil 3.18’de gösterildiği şekilde sivri külah bitişi, katlar ve üst örtü doğru biçimde üretilmiştir.

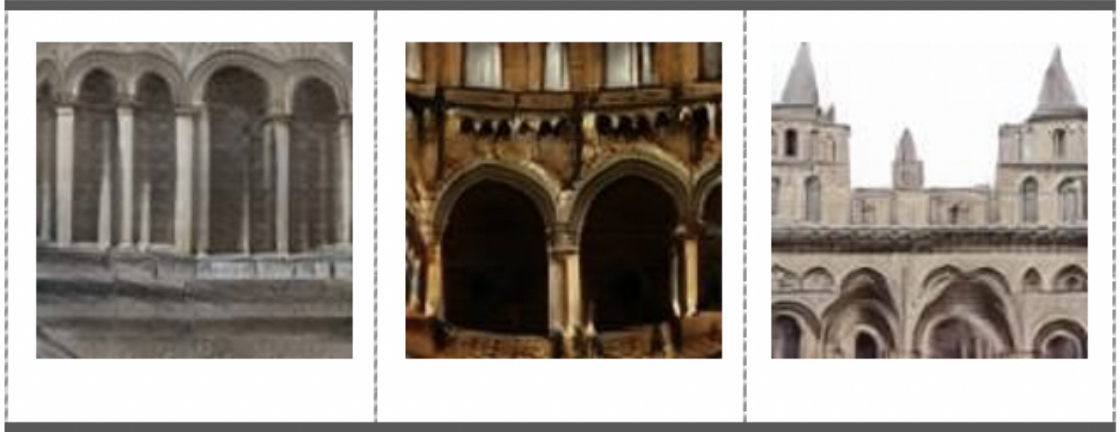
3.3.1.4 İslam Mimarisi Yapı Unsurlarının Değerlendirilmesi

Şekil 3.19’da görüldüğü gibi kapıların simetrik ve daralan kemerleri doğru şekilde üretilmiştir.



Şekil 3.19 Romanesk mimarisi üretilen taç kapı girişi örnekleri

Şekil 3.20’de görüldüğü şekilde Romanesk yapılarda kemer unsuru doğru biçimde oluşturulmamıştır. Simetri özelliği mevcut değildir.



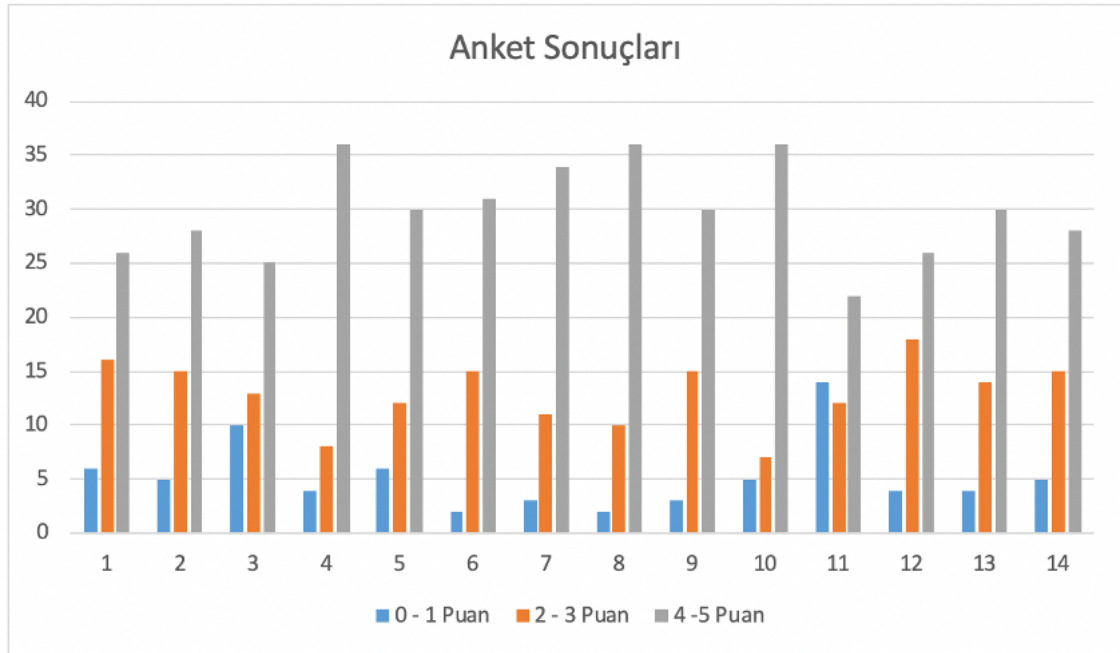
Şekil 3.20 Üretilen romanesk mimari kemer örnekleri

3.4 Üretilen Görüntülerin Gerçekçiliği Üzerine Anket Çalışması

Üretilen görüntülerin gerçekçiliğini değerlendirmek amacıyla bir anket çalışması yapılmıştır. Anket çalışması 48 katılımcı üzerinde gerçekleştirilmiştir. Her biri 4 üretilmiş örnekten oluşan 14 soruda görüntülerin gerçekçi olup olmadığı sorulmuştur. Sorulara ‘gerçekçi değil’ (0) ve ‘gerçekçi’ (5) aralığında puanlar verilmesi istenmiştir. Üretilen görüntü örnekleri en benzer gerçek örneklerle birlikte sunulurken mimari yapıların olması gereken görünümü temsil edilmiştir. Ayrıca katılımcılardan üretilen görüntülerde dikkat çeken ve görüntünün gerçekçiliğini bozan özellikleri belirtmeleri istenmiştir.

12 soruda yer alan 48 üretilmiş görüntüye katılımcıların en az 54%’ü 4 veya 5 puanla oylayarak gerçeğe çok yakın veya gerçekçi olduğunu ifade etmiştir. 3. ve 11.

soruda yer alan 8 görüntüyü ise katılımcıların en az 21%'i 0 veya 1 puanla oylayarak örneklerin gerçekçi olmadığını ifade etmişlerdir. Anketin ayrıntıları Şekil 3.21'de yer almaktadır.



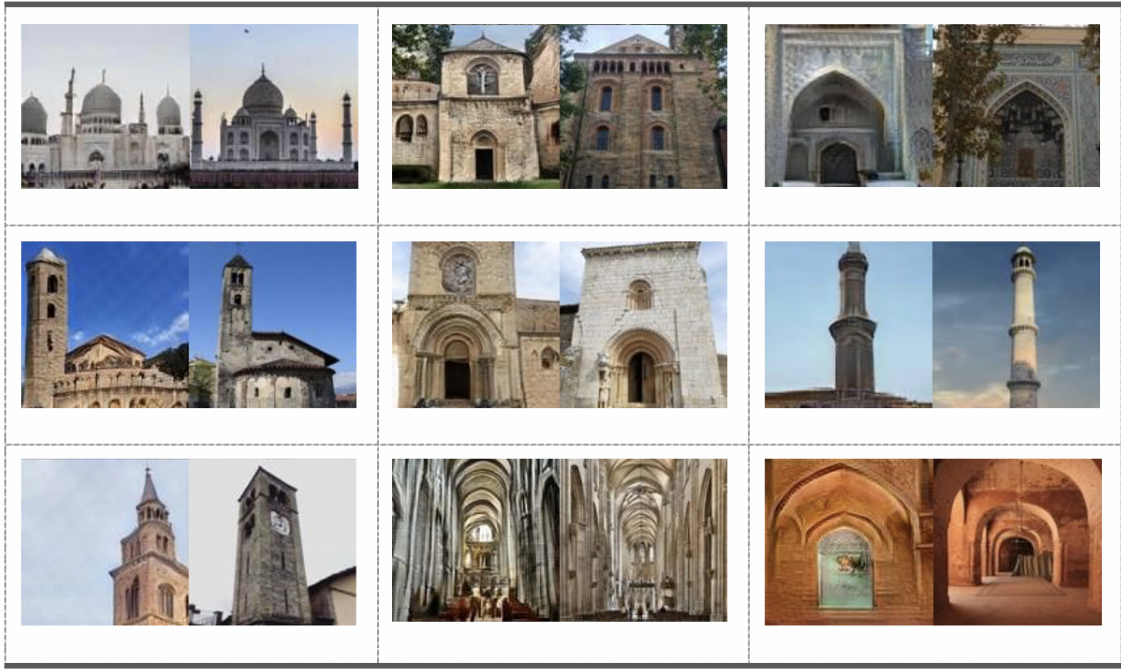
Şekil 3.21 14 sorudan oluşan ankette katılımcıların görüntülere verdiği gerçekçilik puanları

Katılımcılara göre üretilen görüntülerin gerçekçi olmayan özellikleri dört ana problemten oluşmaktadır:

- Görüntülerde keskinliğin yetersiz olması.
- Örneklerde çizgisel kırılma ve kaymaların olması.
- Örneklerin detaylarının yetersiz olması.
- Mimari unsurlarda simetrinin olmaması.

Şekil 3.22'de anket çalışmasında katılımcılar tarafından 'gerçekçi' veya 'gerçeğe çok yakın' olarak ifade edilen görüntülerden birtakım örnek yer almaktadır. Anketin içeriğine dair daha detaylı bilgiler Ekler bölümünde yer almaktadır.

Şekil 3.23'de katılımcılar tarafından gerçekçi bulunmayan görüntüler yer almaktadır. Anket çalışmasına göre, (e) ve (h) görüntülerinde minarelerin, (f)'de kemerlerin simetrik olmaması gerçekçiliği bozan etmenlerden biridir. Ayrıca, üretilen örneklerin keskin hatlara sahip olmaması gerçekçiliği azaltmaktadır. (a), (b), (c), (d), (e) ve (h)'de minarelerin tek bir doğrultuda olmaması ve çizgisel kaymaların olması



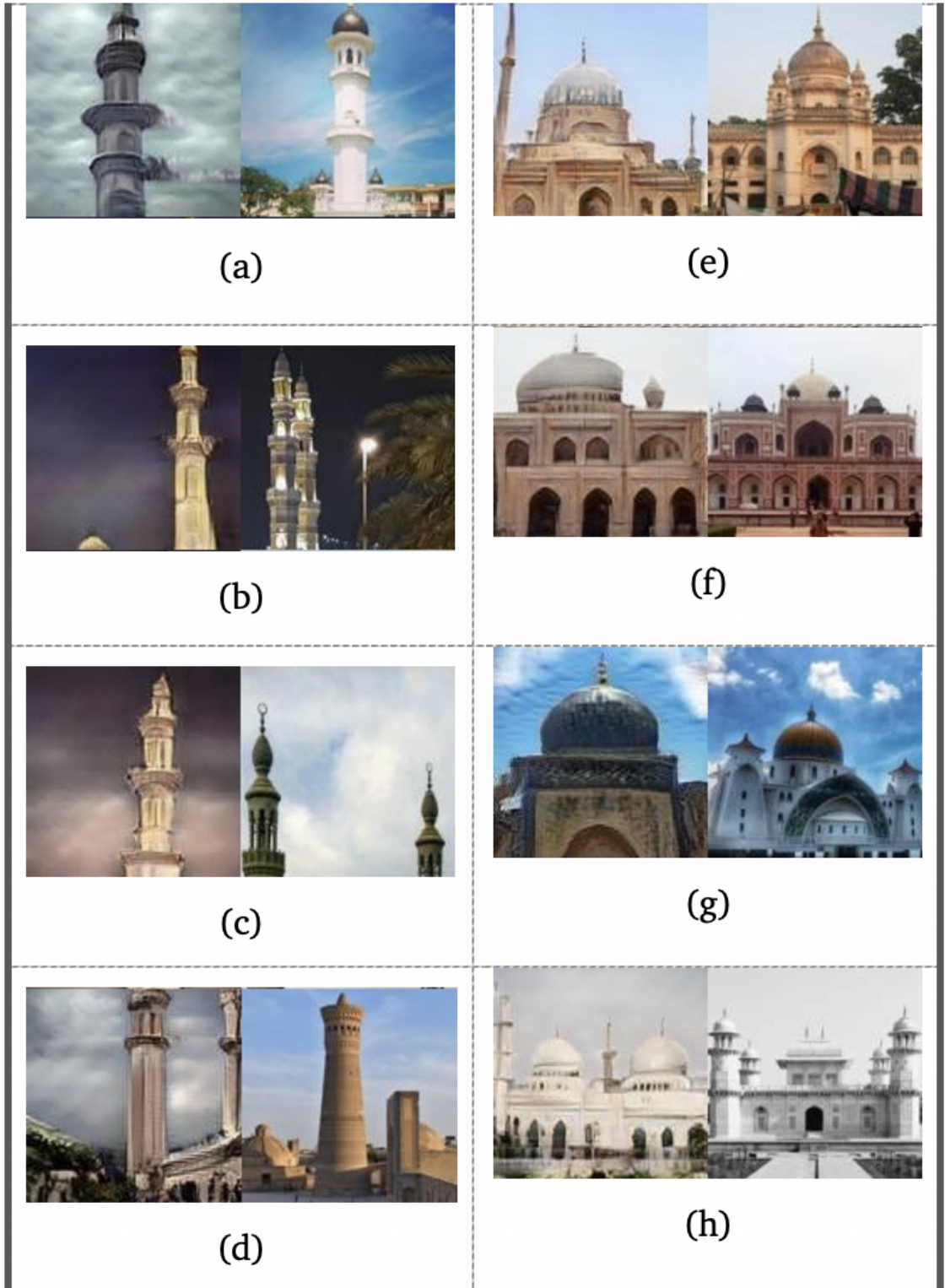
Şekil 3.22 Anket çalışmasında gerçekçi olarak ifade edilen örneklerden bazıları (Görüntü çiftlerinde üretilen görüntü solda ve gerçek görüntü sağda yer almaktadır)

görüntülerin gerçekçiliğini bozmaktadır. Bunun yanında katılımcılar örneklerde detayların yetersiz olduğunu ifade etmişlerdir.

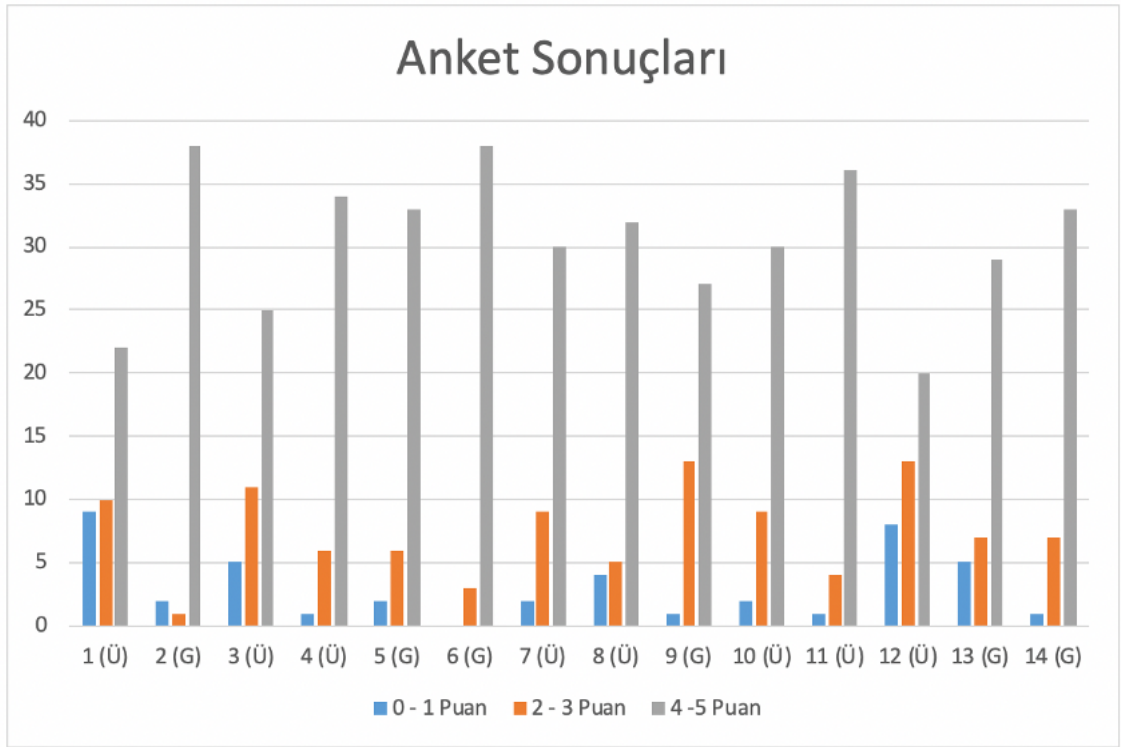
Yapılan diğer bir anket çalışmasında kullanıcılara gerçek ve üretilen görüntüler karıştırılarak verilmiştir. 41 kullanıcıya sunulan ankette her biri 4 görüntüden oluşan 14 soruda görüntüler gerçekçi değil (0) veya gerçekçi (5) olarak 0-5 aralığında oylanmıştır. Soruların 8 tanesinde üretilen örnekler ve 6 tanesinde gerçek örnekler yer almaktadır. Anket sonuçlarına göre üretilen örnekleri içeren 8 sorudan 6'sında katılımcıların en az 61%'i 4-5 puan vermiştir. Üretilen ve gerçek görüntülerin yer aldığı anket sonuçlarına Şekil 3.24'te yer verilmiştir.

Romanesk mimarisi ve İslam mimarisi görüntüleri ayrı ayrı değerlendirildiğinde Romanesk mimarisi örneklerinin daha gerçekçi olduğu görülmektedir. Romanesk mimarisi görüntüleri en az 73% oranında gerçekçi veya gerçeğe yakın olarak oylanmıştır. İslam mimarisi örnekleri ise daha az gerçekçi bulunmuştur.

Şekil 3.25 ve 3.26 incelediğinde Romanesk mimarisi görüntülerinin İslam mimarisi görüntülerine göre daha gerçekçi olduğu gözlenmektedir. Bu durum modelin eğitiminde kullanılan İslam mimarisi örneklerinin çok geniş bir coğrafyada ve büyük bir zaman diliminde yer almasından kaynaklanmaktadır. Bu durum örneklerin çok çeşitlenmesine ve modelin tarzı yansıtan karakteristik özellikleri çıkarmasında zorlanmasına neden olmuştur.



Şekil 3.23 Katılımcılar tarafından gerçekçi bulunmayan görüntüler (Görüntü çiftlerinde üretilen görüntü solda ve gerçek görüntü sağda yer almaktadır)

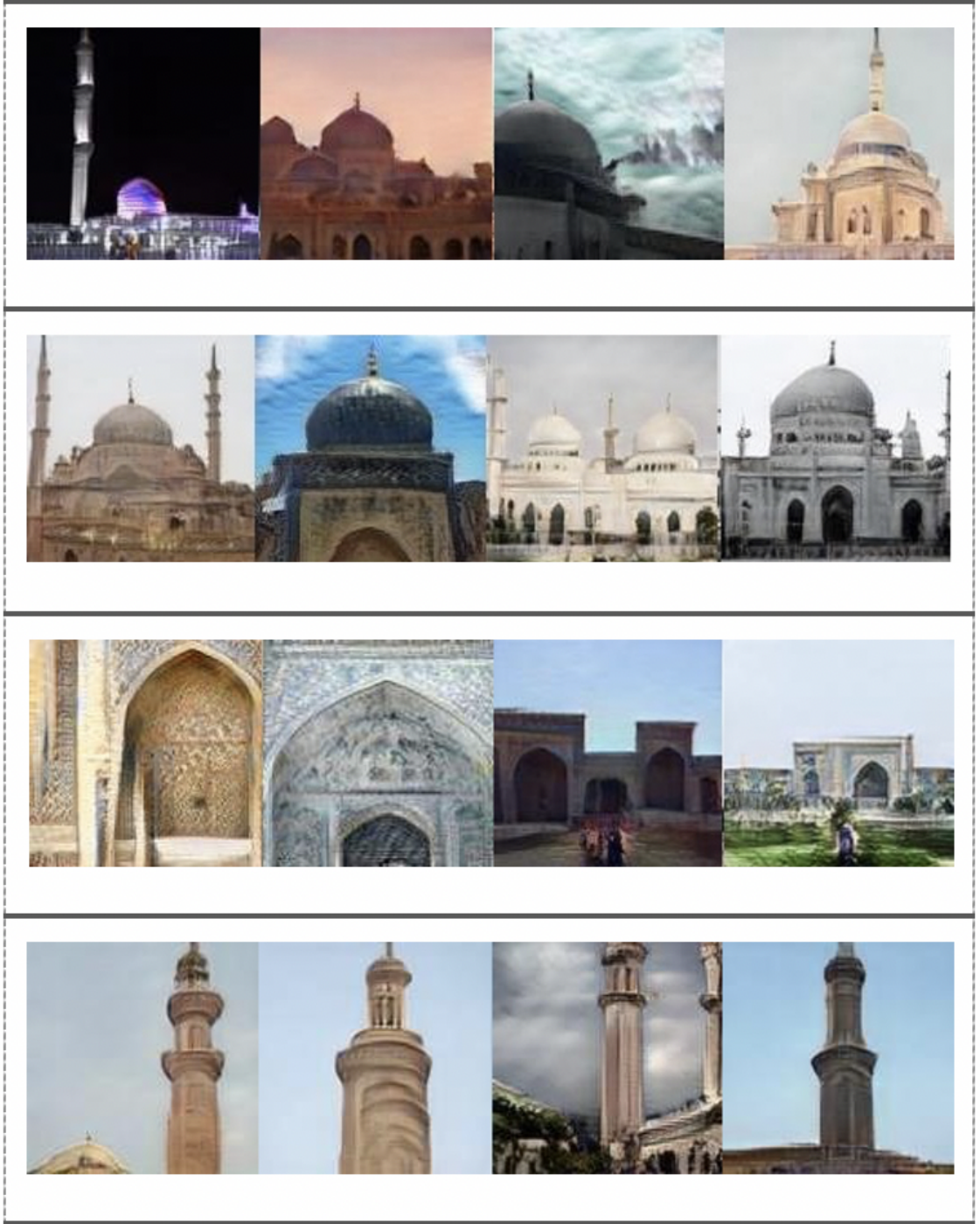


Şekil 3.24 14 sorudan oluşan ankette katılımcıların üretilen ve gerçek görüntülere verdiği gerçekçilik puanları

Önceki paragraflarda değerlendirilen anket çalışması mimari alanında uzman olan 24 katılımcıya da ayrı bir anket olarak yapılmıştır. 8 soruda yer alan üretilen görüntülerde 4 tanesi Romanesk ve 4 tanesi İslam mimarisi örnekleri içermektedir. Bu çalışmada da öncekine benzer şekilde Romanesk mimarisi görüntüleri daha başarılı bulunmuştur. Romanesk örneklerini içeren 4 sorudan 3 tanesi katılımcıların en az 63%'ü tarafından gerçek veya gerçeğe yakın olarak oylanmıştır. Ancak İslam mimarisi örneklerini içeren 4 soruda örneklerin gerçekçi olduğunu düşünenler 42%'yi geçmemektedir. Anketin sonuçlarına dair ayrıntılar Şekil 3.27'de yer almaktadır.

Katılımcılara göre üretilen örneklerde göze çarpan problemler aşağıdaki gibi özetlenebilir:

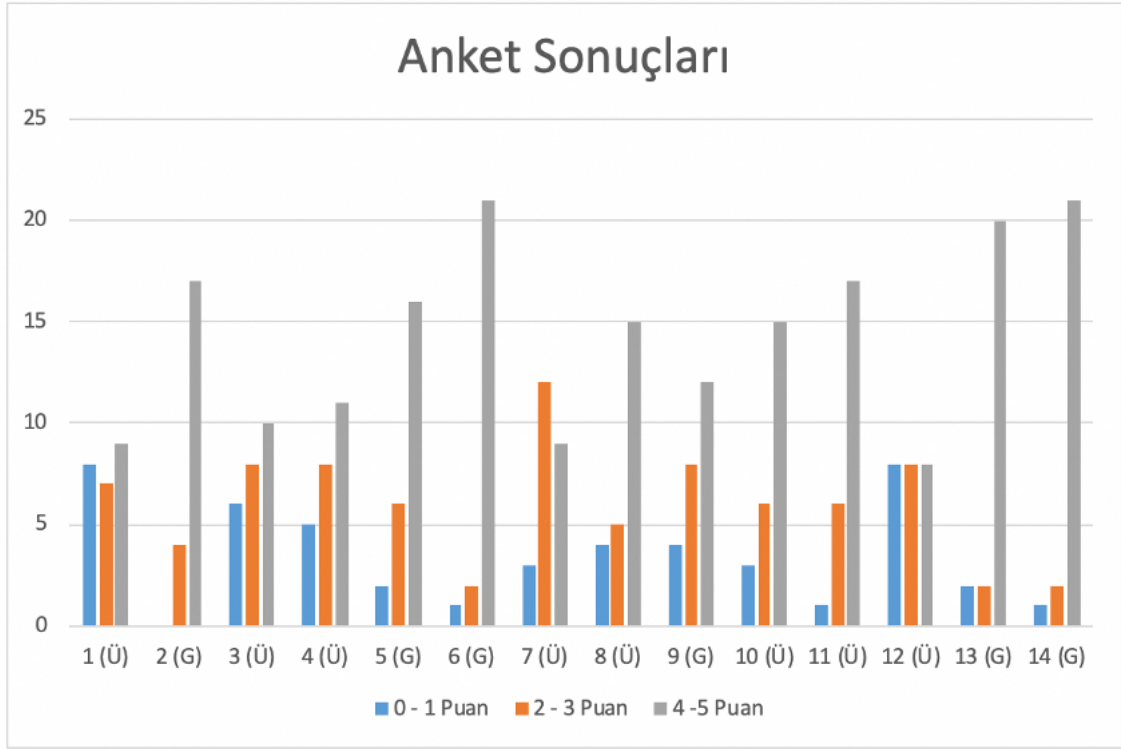
- Yapı unsurlarında aksel kaymalar mevcuttur.
- Görüntülerde bulanıklık mevcuttur.
- Yapı parçalarının birbiriyle ilişkisinde ve perspektif açısından ortantısızlıklar vardır.
- İslam mimarisi çok geniş bir kavramı tanımlıyor o nedenle temsil ettiği yapıların gösterdiği farklı üslupların bir araya gelişi bazı örneklerde tutarlı değildir.



Şekil 3.25 Katılımcıların (sırasıyla) 54%, 61%, 73% ve 49%'u tarafından gerçekçi veya gerçeğe yakın olarak oylananan İslam mimarisi örnekleri



Şekil 3.26 Katılımcıların (sırasıyla) 73%, 78%, 83% ve 88%'i tarafından gerçekçi veya gerçeğe yakın olarak oylanana İslam mimarisi örnekleri



Şekil 3.27 14 sorudan oluşan ankette mimaride uzman katılımcıların üretilen ve gerçek görüntülere verdiği gerçekçilik puanları

- Yapının parçaları doğru yerlere yerleştirilmemiş.

Bu bölümde İslam mimarisine ve Romanesk mimarisine ait üretilen görüntülerin gerçekçiliği üç ayrı anket çalışmasıyla değerlendirilmiştir. Yapılarda gözlemlenen temel problemler açıklanmıştır.

3.5 Modelin Ezberleme Probleminin Değerlendirilmesi

ÜÇA ile daha gerçekçi ve yüksek çözünürlüklü örnekler üretme üzerine çok sayıda çalışma yapılmasına rağmen, ağ eğitirken ağın nasıl genelleştirme yaptığı veya ağın eğitim veri kümesini ezberleme ihtimali konularında yapılan araştırmalar sınırlıdır. ÜÇA'nın yaratıcısı olan Ian Goodfellow [36] çalışmasında bu konuyla ilgili teorik bir yaklaşım geliştirmiştir.

ÜÇA'yı eğitim kümesi örneklerinin kopyasını üreten, ezberleme yapan bir ağ olarak tasarlamıştır. Bunun için üretici ağı sınırlı sayıda gürültü vektörü girdisiyle beslemiştir. Eğitimden sonra ağa yeni bir gürültü verilerek ezberleme test edilmiştir. [37] çalışması üretici ağda ezberleme olasılığının oldukça düşük olduğunu, öğrenmenin ezberlemeye göre çok daha kolay gerçekleştiğini teorik olarak göstermiştir.

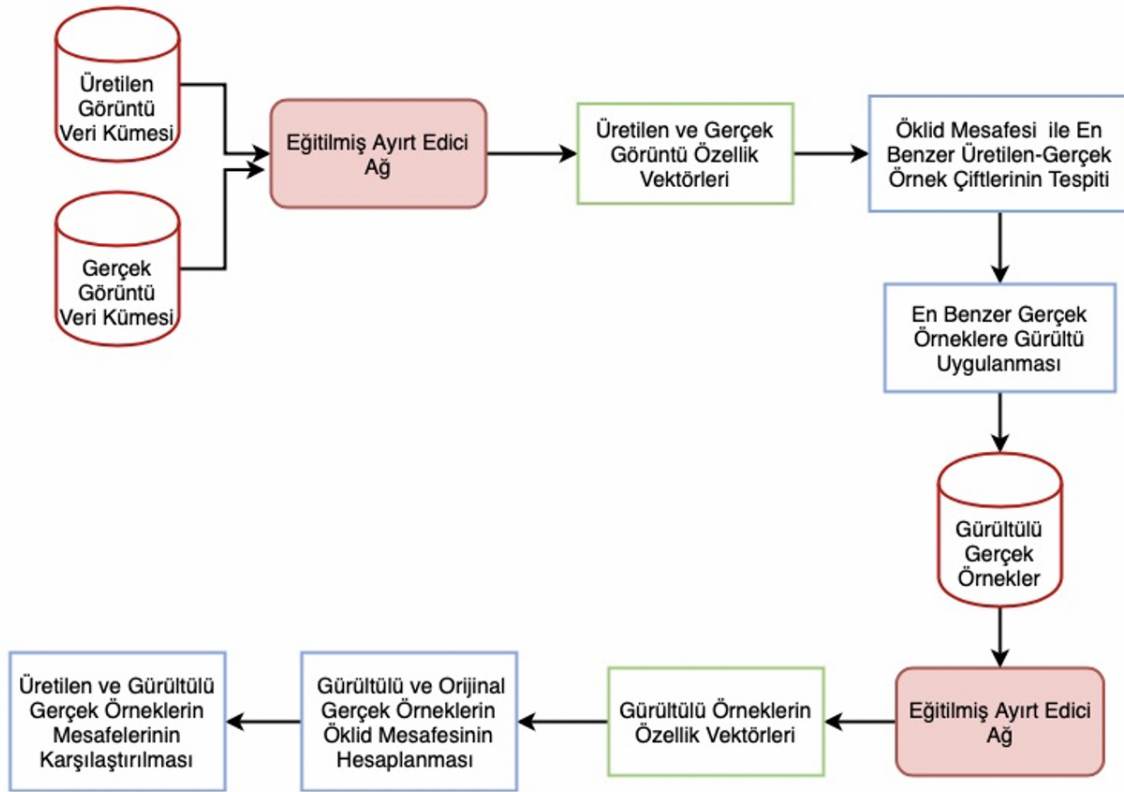
[32] çalışmasında ise ağ değerlendirme metriklerinin ezberlemeyi tespit edip edemeyeceği incelenmiştir. Başlangıç Skoru yöntemi sadece üretilen örnekler üzerinde yapıldığı için, ezberlemeyi tespit edememektedir. Fréchet Başlangıç Mesafesi yöntemi ise ezberlemeyi kısmen tespit ettiği için güvenilir bir yöntem olarak kabul edilmemektedir.

Bu çalışmada, ezberlemeyi tespit etmek için içerik tabanlı görüntü erişimi yöntemi kullanılmıştır. Eğitimi tamamlanmış ayırt edici ağ yardımıyla üretilmiş görüntülerin ve transformasyon uygulanmış görüntülerin gerçek görüntülerle benzerliği hesaplanmıştır. Transformasyon uygulanan görüntülerin gerçek görüntülere daha benzer olması modelin ezberleme olasılığının düşük olduğunu göstermektedir.

Modelin ezberleme yapma olasılığını değerlendirmek için üretilen görüntülere en benzer gerçek görüntüler elde edilmiştir. Elde edilen gerçek görüntülere çeşitli transformasyon işlemleri uygulanarak orijinal gerçek görüntülerle benzerlikleri ölçülmüştür. Üretilen görüntülerle karşılaştırıldığında, transformasyon uygulanan görüntülerin gerçek görüntülere daha benzer olması modelin ezberleme yapma olasılığını azaltmaktadır. Şekil 3.24'te özetlenen ezberleme problemini tespit etme için uygulanan adımlar aşağıda sıralanmıştır:

- Üretilen ve gerçek görüntü örnekleri eğitilmiş ayırt edici ağa girdi olarak verilir ve ağın son katmanından önce 2048 uzunluğundaki özellik vektörleri kaydedilir.
- Üretilen ve gerçek görüntüler arasındaki özellik vektörlerinin benzerliği 2048 eleman çifti ile toplam Öklid Mesafesi hesaplanır.
- En benzer gerçek görüntülere Laplacian filtresi, Gaussian gürültüsü, rasgele karartma, tuz-biber gürültüsü veya kırpma işlemleri uygulanarak transformasyon yapılır.
- Gürültülü görüntü örnekleri ayırt edici ağdan geçirilerek özellik vektörleri elde edilir.
- Gürültülü ve orijinal örneklerin özellik vektörleri arasındaki Öklid mesafesi hesaplanır.
- Üretilen ve gürültülü görüntülerin gerçek görüntülerle olan benzerlikleri karşılaştırılır.

Şekil 3.28 ve 3.29'da gürültü eklenmiş gerçek görüntü örnekleri ve bu görüntülerin orijinal görüntüyle benzerlik mesafesi verilmiştir. Gerçek örneklere Laplacian filtresi,

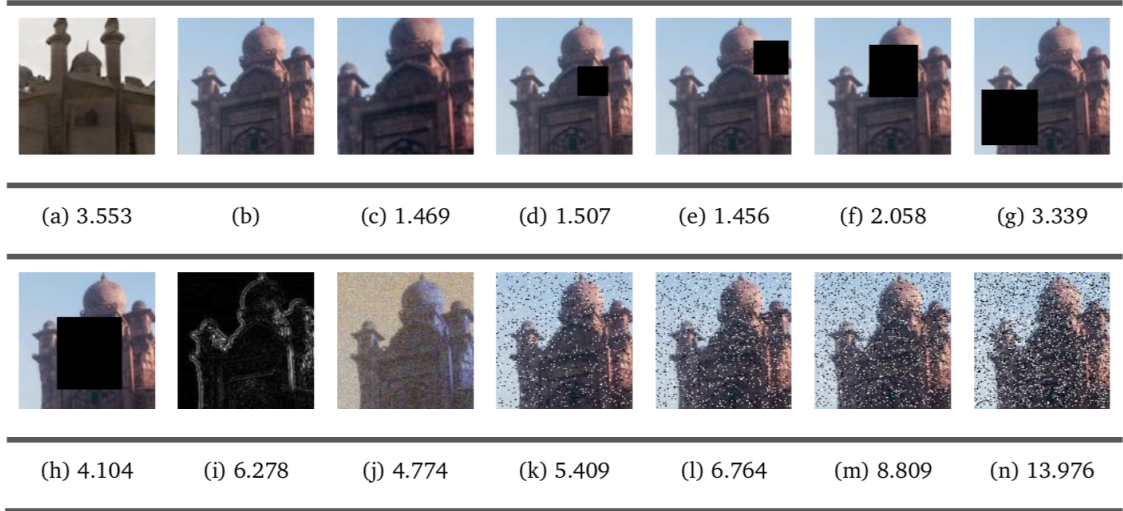


Şekil 3.28 Ezberleme problemini tespit etmede uygulanan adımlar

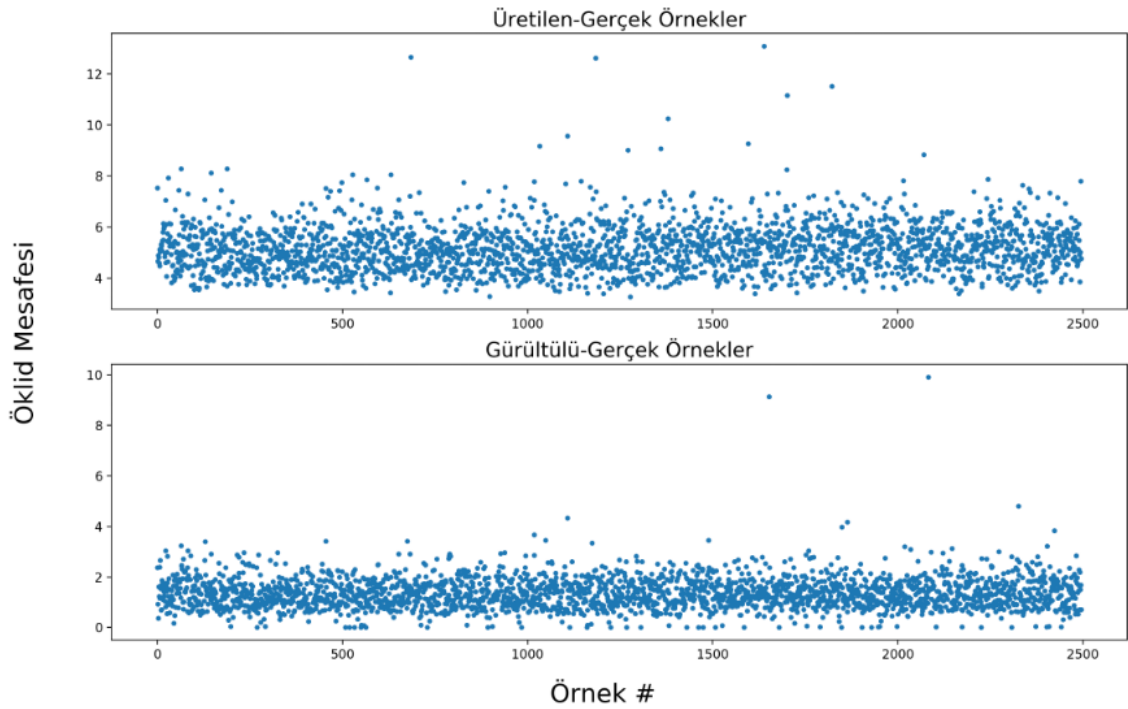
(a) 5.008	(b)	(c) 1.381	(d) 2.169	(e) 1.294	(f) 2.247	(g) 2.188
(h) 3.379	(i) 5.963	(j) 3.514	(k) 4.897	(l) 6.807	(m) 8.633	(n) 14.798

Şekil 3.29 Üretilen (a), en benzer gerçek (b) görüntü, kırpma-yeniden boyutlandırma (c), rasgele karartma (d, e, f, g, h), Laplacian filtresi (i), Gaussian gürültüsü (j) veya tuz-biber gürültüsü (k, l, m, n) uygulanmış görüntüler ve orijinal gerçek görüntüyle benzerlik mesafeleri

kırpma-yeniden boyutlandırma, rasgele karartma, Gaussian gürültüsü veya tuz-biber gürültüsü eklenmiştir. Rasgele karartma görüntü boyutunun 1/25 – 1/4 aralığındaki oranlarında uygulanmıştır. Tuz-biber gürültüsü 0.003 – 0.08 aralığındaki miktarlarda uygulanmıştır.

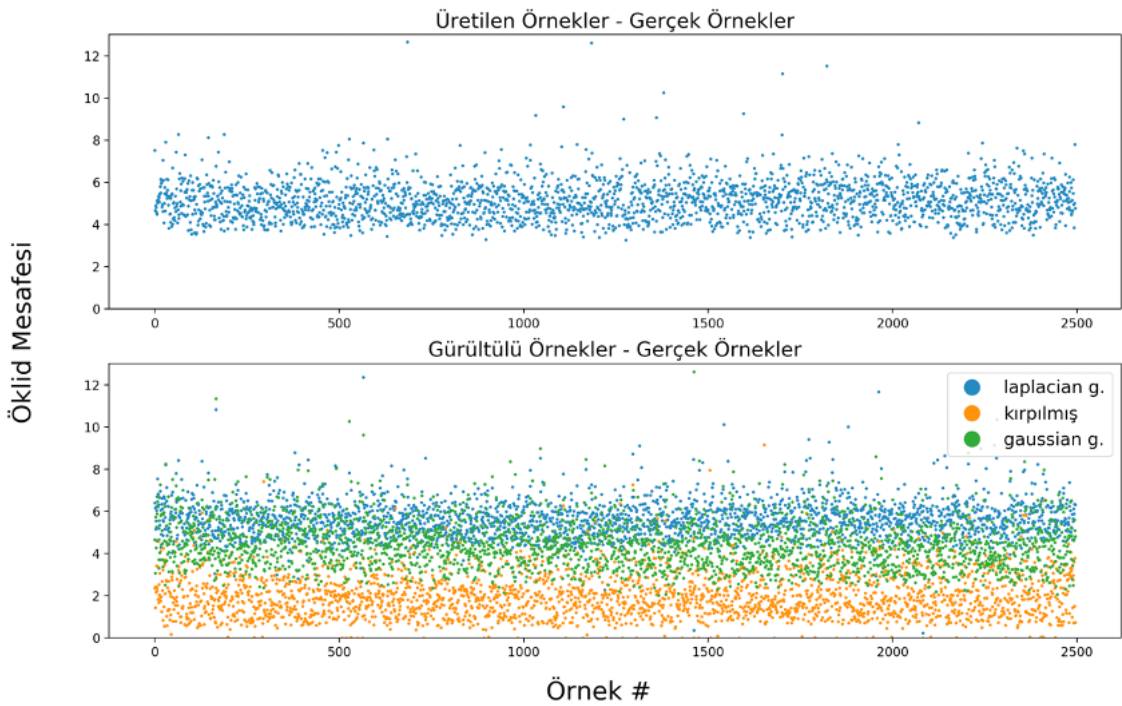


Şekil 3.30 Üretilen (a), en benzer gerçek (b) görüntü, kırpma-yeniden boyutlandırma (c), rasgele karartma (d, e, f, g, h), Laplacian filtresi (i), Gaussian gürültüsü (j) veya tuz-biber gürültüsü (k, l, m, n) uygulanmış görüntüler ve orijinal gerçek görüntüyle benzerlik mesafeleri



Şekil 3.31 Üretilen örneklere en benzer gerçek örneklerin Öklid mesafeleri(üstte), çeşitli gürültüler uygulanmış gerçek örneklerin orijinal örneklere minimum Öklid mesafeleri(alta)

Şekil 3.31’de 1250 İslam ve 1250 Romanesk olmak üzere 2500 üretilen örneği içeren iki grafik mevcuttur. Üstteki grafikte üretilen görüntü örneği ile o örneğe en çok benzeyen gerçek görüntü örneği arasındaki mesafe gösterilmiştir. Üretilen görüntü örnekleri ile gerçek görüntü örnekleri arasındaki mesafe ortalaması 5.17 ve standart sapma değeri 0.93’tür. İkinci grafikte ise en benzer gerçek görüntülere çeşitli gürültüler uygulanmış ve orijinal görüntü örnekleriyle arasındaki minimum benzerlik mesafesi hesaplanmıştır. Bu değerlerin ortalaması 1.36 ve standart sapma değeri 0.63’tür. Gürültülü gerçek örneklerle kıyasla, üretilmiş örneklerin gerçek örneklerle daha az benzediği görülmektedir.



Şekil 3.32 Üretilen görüntülerin gerçek görüntülerle benzerlik mesafeleri (üstte) ve rasgele kırılmış, Gaussian gürültüsü eklenmiş ve Laplacian filtresi uygulanan görüntülerin gerçek görüntülerle benzerlik mesafeleri (altta)

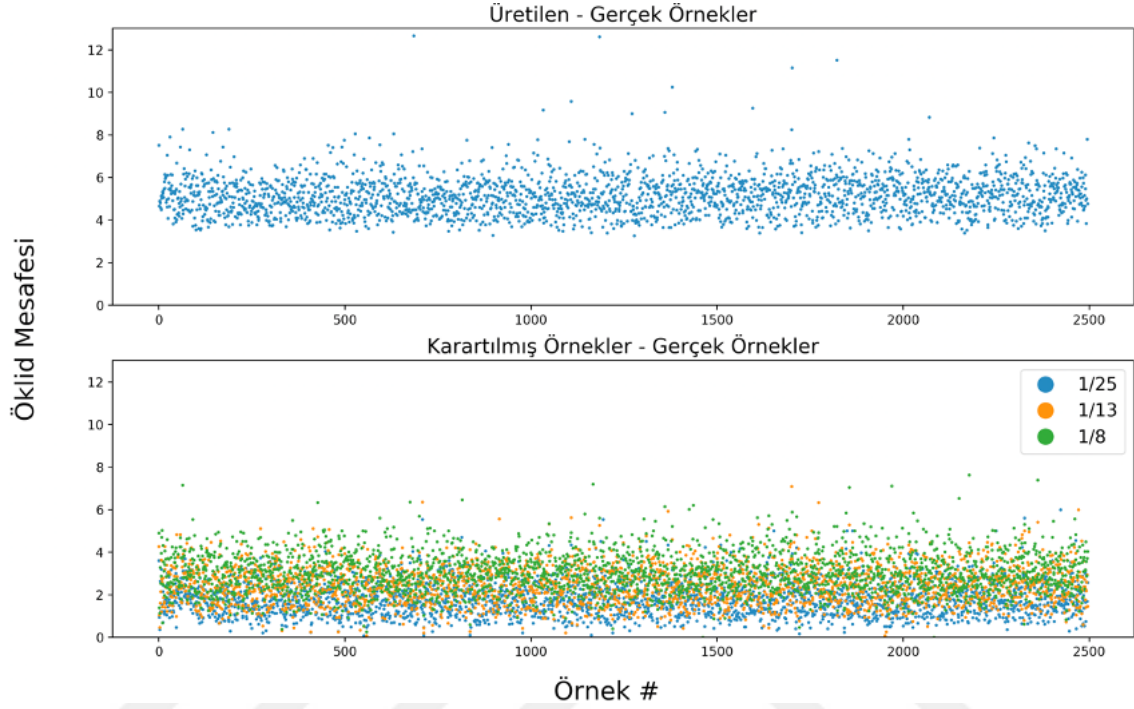
Şekil 3.32’de üst kısımda yer alan grafik üretilen örneklerle bu örneklere en çok benzeyen gerçek görüntülerin Öklid mesafesini göstermektedir. Alt kısımda bulunan grafik ise, Laplacian filtresi, kırpma-yeniden boyutlandırma veya gaussian gürültüsü uygulanmış gerçek görüntülerin orijinal görüntülerle arasındaki Öklid mesafesini göstermektedir.

Laplacian görüntünün kenar özelliklerini ortaya çıkarmada kullanılan bir filtredir. Yalnızca kenar özellikleri kullanılarak benzerlik mesafesi karşılaştırıldığında, üretilen görüntülerin gerçek görüntülere daha benzer olduğu gözlemlenmiştir.

Rasgele kırpma ve yeniden boyutlandırma uygulandığında görüntü ayırt edici özelliklerini koruyarak orijinal görüntüye çok yakın olduğunu gösteren benzerlik

oranları elde edilmiştir.

Gaussian gürültüsü görüntüde rasgele renk dalgalanmalarına neden olmuştur. Gaussian gürültüsü uygulanan görüntülerde örnekleri benzerliği değişkenlik göstermektedir. Ezberlemeyi tespit etmede kullanışlı bir yöntem değildir.

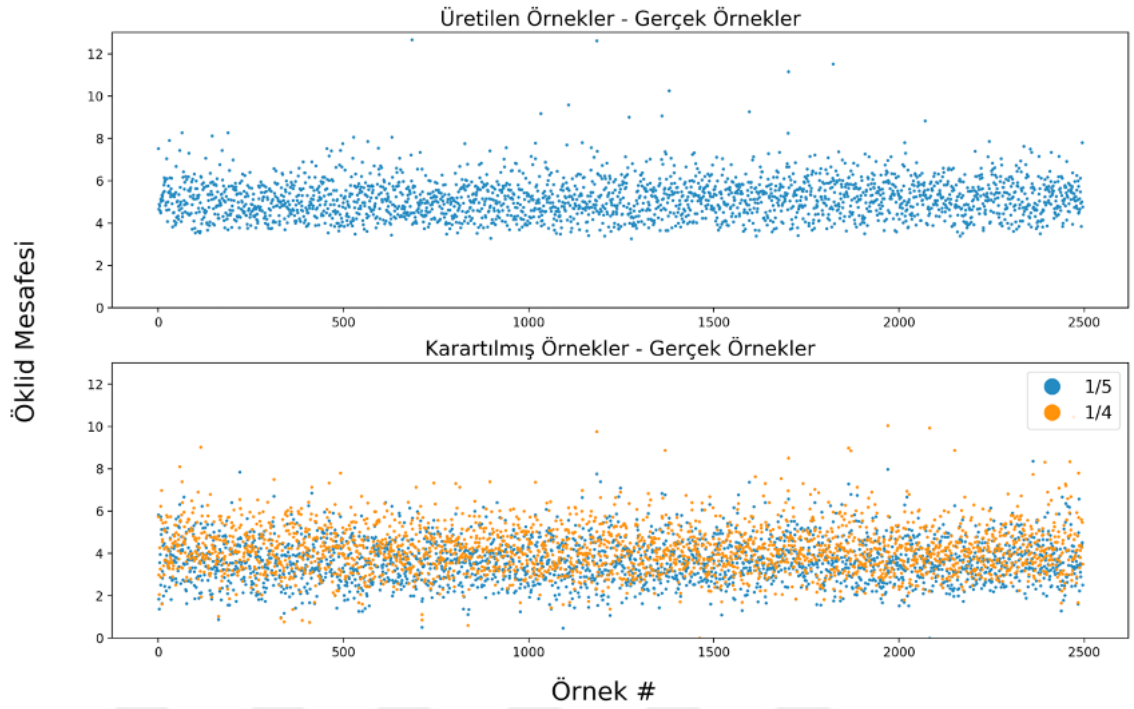


Şekil 3.33 Üretilen görüntülerin gerçek görüntülerle benzerlik mesafeleri (üstte) ve 1/25'i, 1/13'ü ve 1/8'i rasgele karartılan görüntülerin orijinal gerçek görüntülerle benzerlik mesafeleri (altta)

Örneklere görüntü boyutunun 1/25 – 1/4 oranlarında rasgele karartma uygulanmıştır. Şekil 3.33 ve 3.34'a göre üretilen görüntülerle karşılaştırıldığında 1/25, 1/13, 1/8 ve 1/5 oranlarında karartma yapılan gürültülü görüntülerin orijinal örneklerle benzerliği daha fazladır. Ancak 1/4 oranındaki bozulmalar ayırt edici özelliklerin büyük oranda kaybolmasına neden olmuştur. Bu nedenle gürültülü ve orijinal görüntülerin benzerliğini büyük oranda azaltmıştır.

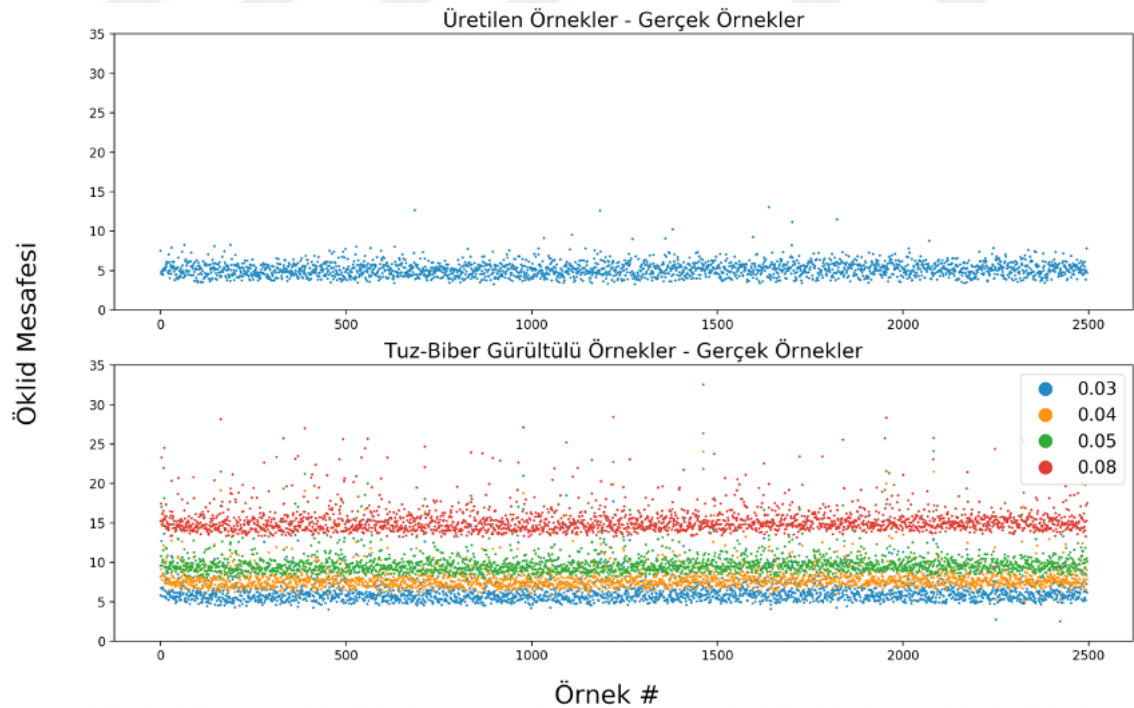
Şekil 3.35'de gösterildiği şekilde görüntülere tuz-biber gürültüsü eklenmesiyle oluşan siyah-beyaz lekeler keskin bir dağılımla görüntülerde büyük değişikliğe neden olmuştur. Tuz-biber gürültüsü örneğin biçimsel ve dokusal özelliklerini yitirmesine sebep olduğu için gürültülü ve orijinal gerçek görüntülerin benzerlikleri düşüktür. Çoğunlukla, üretilen görüntüler gerçek görüntülere daha benzerdir.

Tuz-biber gürültüsü gerçek örneklerin özellikleri üzerinde büyük değişimlere sebep olmuştur. Görüntünün parça-parça ve parça-bütün ilişkisini gözle görülür biçimde bozduğu için ezberlemeyi tespit etmede benzerlik ölçütü olarak kullanılması yanıltıcı

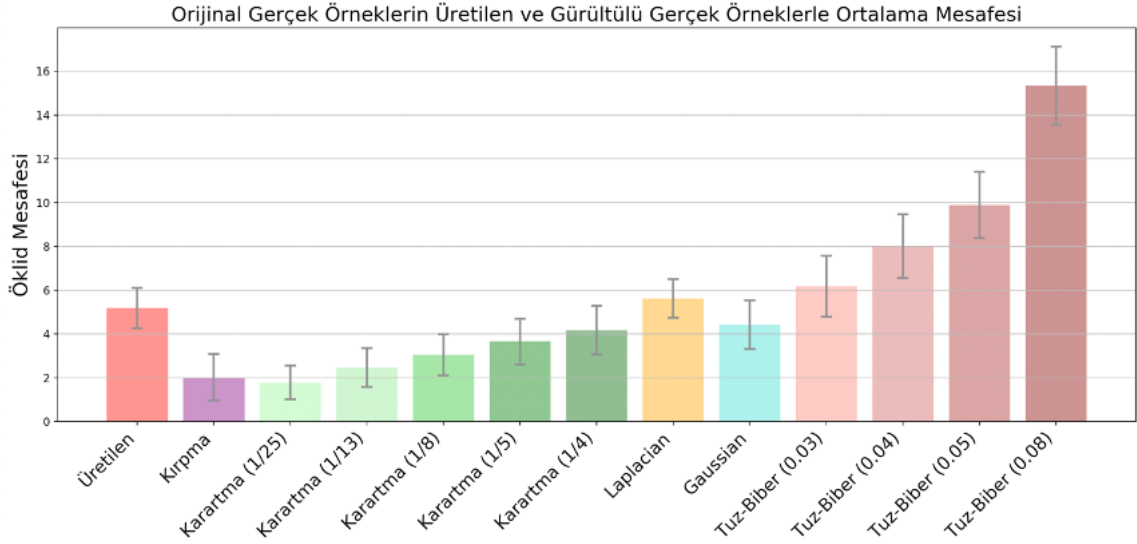


Şekil 3.34 Üretilen görüntülerin gerçek görüntülerle benzerlik mesafeleri (üstte) ve 1/5'i ve 1/4'ü rasgele karartılan görüntülerin orijinal gerçek görüntülerle benzerlik mesafeleri (altta)

olacaktır.



Şekil 3.35 Üretilen görüntülerin gerçek görüntülerle benzerlik mesafeleri (üstte) ve farklı oranlarda tuz-biber gürültüsü eklenmiş görüntülerin orijinal gerçek görüntülerle benzerlik mesafeleri (altta)



Şekil 3.36 Üretilen veya transformasyon uygulanmış görüntülerin orijinal gerçek görüntülerle ortalama mesafesi ve sapma değerleri

Şekil 3.36’de 2500 örnek için, üretilen görüntülerin ve farklı miktarlarda transformasyon uygulanmış gerçek görüntülerin orijinal gerçek görüntülerle olan ortalama Öklid uzaklığı ve standart sapma değerleri verilmiştir.

“Üretilen” sütununda üretilen örnekler ve onlara en benzer gerçek örneklerin ortalama Öklid mesafesi ve sapması görülmektedir. Diğer sütunlarda ise sırasıyla kırpılarak yeniden boyutlandırılmış, 1/25 – 1/4 oranlarında rasgele karartma uygulanmış, Laplacian filtresi uygulanmış, Gaussian gürültüsü uygulanmış ve 0.003 – 0.08 oranlarında tuz biber gürültüsü eklenmiş gerçek görüntülerin orijinal gerçek görüntüye olan uzaklığı görülmektedir.

Üretilen örneklerle karşılaştırıldığında, kırpma-yeniden boyutlandırma, rasgele karartma veya Gaussian gürültüsü uygulanan gerçek örnekler orijinal gerçek örneklerle daha benzerdir. Üretilen örneklerin orijinal gerçek görüntüye daha benzer olduğu örnekler ise Laplacian filtresi veya tuz biber gürültüsü eklenmiş görüntülerdir. Laplacian filtresi uygulanan görüntülerde sadece kenar özellikleri bulunduğu için orijinal görüntüyle benzerliği azdır. Tuz-biber gürültüsü az miktarda uygulandığında dahi görüntüdeki özellikler lokal olarak bozduğu için çalışma prensibi lokal özelliklere dayanan konvolüsyonel ağların benzerliği tespit edememesine neden olmaktadır.

Sonuçta üretilen görüntülerle karşılaştırıldığında, transformasyon uygulanan görüntülerin orijinal gerçek örneklerle daha benzer olması modelin ezberleme olasılığının düşük olduğunu göstermektedir.

4 SONUÇ VE ÖNERİLER

Bir örnek uzayının oluşturduğu olasılık dağılımını öğrenerek bu dağılıma uygun yeni örnekler üretmeyi sağlayan üretici modeller ile video ve görüntü üretme üzerine çok sayıda çalışma yapılmıştır. Bu çalışmada derin öğrenme tabanlı üretici modellerden birisi olan Üretici Çekişmeli Ağlar'ın çeşitli mimari tarzlarda gerçekçi yapı örnekleri üretmedeki başarımları üzerine bir araştırma yapılmıştır. Yapılar ait oldukları mimari tarzların özelliklerine göre birbirinden ayrılrsa da birçok benzer özelliğe de sahiptirler. Ayrıca, mimari yapıların yapısal, dokusal ve detay özellikleri çok çeşitli ve karmaşıktır. Görüntü üretmede kullanılan CIFAR [38], LSUN [39] ve ImageNet [40] gibi nesne veri kümeleri yerine, bu tür karmaşık özelliklere sahip bir mimari yapı veri kümesi üzerinde çalışarak ÜÇA modellerinin başarı sınırlarını, zayıf ve güçlü yönlerini test etmek amaçlanmıştır. Bu çalışma için oluşturulan veri kümesinde Romanesk mimarisine ait yaklaşık 19000 kilise, katedral, kale ve şato görüntü örneği ve İslam mimarisine ait yaklaşık 20000 cami, medrese ve kütüphane görüntü örneği mevcuttur. Bu yapıların genel dış cephe görüntülerinin yanında, minare, çan kulesi, sunak, iç mekan, kemer, kapı girişi, vitray, mukarnas ve eyvan gibi mimari yapı unsurları da veri kümesinde yer almaktadır. Görüntü veri kümesini standart hale getirmek için farklı formattaki tüm görüntüler RGB biçimine dönüştürülmüştür. Görüntüler yüksek çözünürlükte olduğu için boyut azaltma adımıyla bilgi kaybını azaltmak amacıyla tüm görüntüler kırılarak, gereksiz nesne ve arka plan bilgisi atılmıştır. Ayrıca, yüksek çözünürlükteki her bir görüntü için 5 kez rastgele kırma yapılarak veri kümesi genişletilmiştir. Yaklaşık 114000 Romanesk mimarisine ait ve yaklaşık 120000 İslam mimarisine ait görüntü örneği 128x128 boyutuna dönüştürülmüştür.

Bu çalışma için oluşturulan veri kümesinde Romanesk mimarisine ait yaklaşık 19000 kilise, katedral, kale ve şato görüntü örneği ve İslam mimarisine ait yaklaşık 20000 cami, medrese ve kütüphane görüntü örneği mevcuttur. Bu yapıların genel dış cephe görüntülerinin yanında, minare, çan kulesi, sunak, iç mekan, kemer, kapı girişi, vitray, mukarnas ve eyvan gibi mimari yapı unsurları da veri kümesinde yer almaktadır. Görüntü veri kümesini standart hale getirmek için farklı formattaki tüm görüntüler

RGB biçimine dönüştürülmüştür. Görüntüler yüksek çözünürlükte olduğu için boyut azaltma adımı bilgi kaybını azaltmak amacıyla tüm görüntüler kırılarak, gereksiz nesne ve arka plan bilgisi atılmıştır. Ayrıca, yüksek çözünürlükteki her bir görüntü için 5 kez rastgele kırma yapılarak veri kümesi genişletilmiştir. Yaklaşık 114000 Romanesk mimarisine ait ve yaklaşık 120000 İslam mimarisine ait görüntü örneği 128x128 boyutuna dönüştürülmüştür.

Üretici Çekişmeli Ağlar, üretici ve ayırt edici adı verilen iki ayrı çok katmanlı ağdan oluşur. Ayırt edici ağ bir görüntü örneğinin gerçek veya üretilmiş olduğunu anlamaya çalışır. Üretici ağ yeni örnekler üreterek ayırt edici ağı kandırmaya çalışır. Üretici ve ayırt edici ağ birbiriyle yarış halinde olan iki rakip ağdır. Üretici ağ rassal bir gürültü vektörünü girdi olarak alır ve sahte görüntü örneği üretir. İki ayrı girdi ile beslenen ayırt edici ağa gerçek veri kümesinden bir örnek ve üretici ağın yarattığı görüntü örneği verilir. Ayırt edici ağ bu iki örneği de üretilmiş veya gerçek olarak sınıflandırır. Ayırt edici ağın sınıflandırma çıktısına göre ayırt edici ve üretici ağa geri besleme yapmak için iki ayrı kayıp fonksiyonu hesaplanır. Ayırt edici ağda, üretici ağın kaybı üretilmiş örneğe göre, ayırt edici ağın kaybı ise hem üretilmiş hem de gerçek örnek üzerinden hesaplanarak her iki ağın ağırlık parametreleri güncellenir. Bu adımlar diğer örnekler için de benzer şekilde uygulanır. Sonuçta, bir ağın parametre güncellemesi diğer ağın parametre değişimini tetiklemiyorsa eğitim sonlandırılır. Üretici ağın en güncel ağırlık parametreleriyle görüntü örnekleri üretilir.

Bu çalışmada karmaşık ve çok sınıflı veri kümeleri üzerinde başarısını kanıtlamış olan BigGan [15] Üretici Çekişmeli Ağ modelinden yararlanılmıştır. Mimari görüntü örnekleri üretmek için BigGAN modeli üzerinde gerekli düzenlemeler yapılmıştır. BigGan modelinin omurgasını ResNet50 [19] ve SAGAN [14] mimarileri oluşturmaktadır. Ağ derinliğinin artmasıyla ortaya çıkan genel bir sorun olan Gradyan Yok Olması (Vanishing Gradient) problemine çözüm getirdiği için ResNet50 mimarisi, daha derin ağlar oluşturulmasına ve dolayısıyla daha karmaşık veri kümelerinde çalışılmasına olanak sağlamıştır. SAGAN mimarisinin getirdiği bir yenilik olan Öz İlgili Ağlar ise görüntülerde parça-parça ve parça-bütün ilişkisini semantik olarak tespit ederek daha gerçekçi ve detaylı görüntü örnekleri üretmeyi sağlamıştır.

Üretici Çekişmeli Ağlar diğer derin ağlardan farklı olarak, model mimarisinin yanında optimizasyon yöntemlerine ve parametrelerine karşı çok hassas modellerdir. Modele uygulanan Adam optimizasyon algoritması kayıp fonksiyonundan elde edilen gradyanı bir rastgele değişken olarak ele alır. Gradyanın beklenen değeri ve varyansına göre ağırlık parametrelerini günceller. Ağırlık güncellemeleri genelleştirilmiş gradyan üzerinden yapılmış olur. Böylece güncellemelerde çok büyük sapmalar engellenir ve model daha kararlı hale gelir. Bu çalışmada kullanılan modelin genelleştirme

yeteneğini arttırmak için çeşitli regülarizasyon yöntemleri uygulanmıştır. Ağ eğitimi sırasında nöronların parametreleri sık sık değiştiği için, bir katmanda yapılan değişiklik kendinden sonraki katmanları da zincirleme olarak etkilemektedir. Bu durum nöronların girdi dağılımlarının sürekli değişmesine yol açmaktadır. Sonuçta, daha kararsız hale gelen modelin eğitimi zorlaşmaktadır. Bu problemi çözmek için nöronların aktivasyon çıktıları Grup Normalizasyonu ile normalize edilmiştir. Ayrıca, üretici ağı daha küçük (0.0001) ve ayırt edici ağı daha büyük (0.0004) öğrenme oranları verilmiştir. İki Zamanlı Güncelleme Kuralı adı verilen bu yöntem ile üretici ağ ayırt edici ağı kandırmak için daha küçük ve temkinli adımlarla ilerlemektedir. Bunun yanında yaklaşık 39000 veri kümesi örneği rastgele kırılarak 234000 örneğe çıkarılmıştır. Veri kümesinin genişletilmesi de yine bir regülarizasyon yöntemidir.

Bu çalışmada modelin başarısını değerlendirmek için Başlangıç Skoru ve Fréchet Başlangıç Mesafesi ölçütlerinden yararlanılmıştır. Başlangıç Skoru çok sınıflı örnek uzaylarında, aynı sınıf örneklerinin birbirine yakın ve homojen bir dağılıma sahip olması gerektiğini, farklı sınıfların dağılımlarının ise birbirinden uzak olması gerektiğini söyler. Başlangıç Skoru sadece üretilen görüntü örneklerini ele aldığı için performans metriği olarak yeterli değildir. Bu nedenle üretilen görüntülerin kalitesini ölçmek için Fréchet Başlangıç Mesafesi ölçütünden de yararlanılmıştır. Fréchet Başlangıç Mesafesi üretilmiş ve gerçek görüntü örneklerinin dağılımları arasındaki uzaklığı ölçer. Özetle, FBM değerinin düşük; BS değerinin yüksek olması modelin başarılı olduğunu gösterir.

Modelin başarısı FBM ve BS performans metrikleriyle, mimari özellikler açısından ve ezberleme problemi açısından değerlendirilmiştir.

Diğer çalışmalarla kıyaslandığında, bu çalışmada elde edilen 20 FBM değeri üretilen görüntülerin kaliteli ve gerçekçi olduğu kanıtlamıştır. Ancak BS değerinin 4 olması üretilen görüntü örneklerinde çeşitliliğin az olduğunu göstermiştir. Bu durum eğitim esnasında seçilen mini-küme büyüklüğünün düşük olmasından kaynaklanıyor olabilir.

Üretilen Romanesk ve İslam mimarisi görüntü örnekleri yapısal, dokusal ve datay özellikler açısından değerlendirilmiştir. Geometrik özellikleriyle ve bir bütün olarak mimari örneklerini üretmede model yapısal açıdan başarılıdır. Model doku özelliklerini oluşturmada kısmen zorlanmasına rağmen, detayları üretmede yetersizdir. Bu durum, eğitilen görüntülerin boyutunun düşük olmasından kaynaklanabilir. Detaylar mimari yapıların tarzlarını belirlemede önemli özellikler olduğu için görüntüler orijinal boyutuyla ve parçalar halinde eğitilip sonradan birleştirilirse çok daha başarılı mimari görüntü örnekleri üretilibicektir.

Üretilen görüntülerin gerçekçiliğini nicel yöntemlerle tespit etmek için 48 katılımcıyla

bir anket düzenlenmiştir. Her biri 4 üretilmiş ve 4 gerçek görüntü içeren 14 soruluk ankette üretilen görüntülerin gerçekçiliği 0-5 aralığında oylanmıştır. 12 sorudaki 48 görüntü için katılımcıların en az 54%'ü gerçekçi veya gerçeğe yakın oyu kullanmıştır. Gerçekçi bulunmayan 8 örnek için katılımcılar örneklerdeki temel problemleri keskinliğin düşük olması, çizgisel kaymalar ve detaylandırmanın eksikliği olarak yorumlamışlardır.

Bunun yanında görüntülerin kalitesi mimaride uzman bir akademisyen tarafından da teknik olarak değerlendirilmiştir. İslam mimarisi örnekleriyle karşılaştırıldığında model Romanesk mimarisi görüntü örnekleri üretmede daha başarılıdır. Bunun temel sebebi İslam mimarisinin Avrupa, Afrika ve Asya'yı kapsayan geniş bir coğrafyada farklı malzeme, doku ve biçim özelliklerinde çok çeşitli örneklerinin bulunmasıdır. Ayrıca İslam mimarisi yedinci yüzyıldan günümüze kadar olan uzun bir zaman diliminde değişime uğrayan bir üsluptur. Romanesk mimarisi ise yalnızca Doğu Avrupa ve 4 yüzyıllık zaman dilimi ile sınırlı, karakteristik özelliklere sahip olan bir alt üslup niteliğindedir. İslam mimarisi örneklerinde çeşitliliğin çok olması modeli karmaşıktırmaktadır. Bu nedenle Romanesk mimarisi örnekleriyle karşılaştırıldığında İslam mimarisi örnekleri daha başarısızdır.

Diğer modellerin örnekleri ile kıyaslandığında bu çalışmada üretilen görüntülerin daha kaliteli ve gerçekçi olması ÜÇA modelinin gerçek veriyi ezberleme olasılığını düşündürmüştür. Ezberlemeyi tespit etmede BS ve FBM yöntemleri yeterli olmadığı için İçerik Tabanlı Görüntü Erişimi yöntemi yardımıyla eğitilmiş ayırt edici ağ kullanılarak yeni bir yöntem geliştirilmiştir. Ağa girdi olarak verilmiş örneklerin son katmandan önce elde edilen özellik vektörleriyle Öklid mesafesi hesaplanarak üretilen görüntülere en benzer gerçek görüntüler ve benzerlik mesafeleri hesaplanmıştır. En benzer gerçek görüntülere çeşitli gürültü ve filtrelerle transformasyon uygulandıktan sonra orijinal biçimleriyle Öklid mesafeleri hesaplanmıştır. Üretilen örneklerle karşılaştırıldığında, transformasyon uygulanmış görüntülerin orijinal görüntülere daha benzer olması modelin ezberleme yapmış olma olasılığının düşük olduğunu göstermiştir.

- [1] I. Goodfellow, "Nips 2016 tutorial: Generative adversarial networks," *arXiv preprint arXiv:1701.00160*, 2016.
- [2] A. V. Oord, N. Kalchbrenner, K. Kavukcuoglu, "Pixel recurrent neural networks," M. F. Balcan, K. Q. Weinberger, Eds., ser. *Proceedings of Machine Learning Research*, vol. 48, New York, New York, USA: PMLR, 20–22 Jun 2016, pp. 1747–1756. [Online]. Available: <http://proceedings.mlr.press/v48/oord16.html>.
- [3] A. van den Oord, N. Kalchbrenner, O. Vinyals, L. Espeholt, A. Graves, K. Kavukcuoglu, "Conditional image generation with pixelcnn decoders," in *NIPS*, 2016.
- [4] D. Alexandre, C. Chang, W. Peng, H. Hang, "An autoencoder-based learned image compressor: Description of challenge proposal by NCTU," *CoRR*, vol. abs/1902.07385, 2019. arXiv: 1902.07385. [Online]. Available: <http://arxiv.org/abs/1902.07385>.
- [5] P. Akyazi, T. Ebrahimi, "Learning-based image compression using convolutional autoencoder and wavelet decomposition," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019.
- [6] S. Zebang, K. Sei-ichiro, "Densely connected autoencoders for image compression," in *Proceedings of the 2nd International Conference on Image and Graphics Processing*, ser. ICIGP '19, Singapore, Singapore: Association for Computing Machinery, 2019, pp. 78–83, ISBN: 9781450360920. DOI: 10.1145/3313950.3313965. [Online]. Available: <https://doi.org/10.1145/3313950.3313965>.
- [7] R. Zhang, P. Isola, A. A. Efros, "Colorful image colorization," in *European conference on computer vision*, Springer, 2016, pp. 649–666.
- [8] A. Dosovitskiy, T. Brox, "Generating images with perceptual similarity metrics based on deep networks," in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, R. Garnett, Eds., Curran Associates, Inc., 2016, pp. 658–666. [Online]. Available: <http://papers.nips.cc/paper/6158-generating-images-with-perceptual-similarity-metrics-based-on-deep-networks.pdf>.
- [9] X. Chen, D. P. Kingma, T. Salimans, Y. Duan, P. Dhariwal, J. Schulman, I. Sutskever, P. Abbeel, "Variational lossy autoencoder," *arXiv preprint arXiv:1611.02731*, 2016.
- [10] D. R. Ha, D. Eck, "A neural representation of sketch drawings," *ArXiv*, vol. abs/1704.03477, 2018.

- [11] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, "Generative adversarial nets," in *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, ser. NIPS'14, Montreal, Canada: MIT Press, 2014, pp. 2672–2680.
- [12] A. Radford, L. Metz, S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," Nov. 2015.
- [13] T. Karras, T. Aila, S. Laine, J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," *arXiv preprint arXiv:1710.10196*, 2017.
- [14] H. Zhang, I. Goodfellow, D. Metaxas, A. Odena, "Self-attention generative adversarial networks," in *International Conference on Machine Learning*, 2019, pp. 7354–7363.
- [15] A. Brock, J. Donahue, K. Simonyan, "Large scale gan training for high fidelity natural image synthesis," *arXiv preprint arXiv:1809.11096*, 2018.
- [16] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.
- [17] S. Chaillou, *Archigan: A generative stack for apartment building design*, 2019.
- [18] D. Bau, J.-Y. Zhu, H. Strobelt, B. Zhou, J. Tenenbaum, W. Freeman, A. Torralba, "Visualizing and understanding generative adversarial networks (extended abstract)," Jan. 2019.
- [19] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [20] M. LLC. (2019). Detailed guide to understand and implement resnets, [Online]. Available: <https://cv-tricks.com/keras/understand-implement-resnets> (visited on 05/20/2020).
- [21] M. Toğaçar, B. Ergen, "Biyomedikal görüntülerde derin öğrenme ile mevcut yöntemlerin kıyaslanması," *Firat Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 31, pp. 109–121, 2019, ISSN: 1308-9072.
- [22] M. Sanjeevi. (2019). Ch:14 generative adversarial networks (gan's) with math, [Online]. Available: <https://medium.com/deep-math-machine-learning-ai/ch-14-general-adversarial-networks-gans-with-math-1318faf46b43> (visited on 05/20/2020).
- [23] M. Arjovsky, S. Chintala, L. Bottou, "Wasserstein generative adversarial networks," D. Precup, Y. W. Teh, Eds., ser. Proceedings of Machine Learning Research, vol. 70, International Convention Centre, Sydney, Australia: PMLR, Jun. 2017, pp. 214–223. [Online]. Available: <http://proceedings.mlr.press/v70/arjovsky17a.html>.
- [24] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang, S. Paul Smolley, "Least squares generative adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2794–2802.
- [25] H.-W. Dong, Y.-H. Yang, "Towards a deeper understanding of adversarial losses," *arXiv preprint arXiv:1901.08753*, 2019.

- [26] D. Kingma, J. Ba, “Adam: A method for stochastic optimization,” *International Conference on Learning Representations*, Dec. 2014.
- [27] M. Mirza, S. Osindero, “Conditional generative adversarial nets. arxiv 2014,” *arXiv preprint arXiv:1411.1784*, 2014.
- [28] Y. Wu, K. He, “Group normalization,” in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 3–19.
- [29] K. Kurach, M. Lučić, X. Zhai, M. Michalski, S. Gelly, “A large-scale study on regularization and normalization in gans,” in *International Conference on Machine Learning*, 2019, pp. 3581–3590.
- [30] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, S. Hochreiter, “Gans trained by a two time-scale update rule converge to a local nash equilibrium,” in *Advances in neural information processing systems*, 2017, pp. 6626–6637.
- [31] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, X. Chen, “Improved techniques for training gans. arxiv 2016,” *arXiv preprint arXiv:1606.03498*,
- [32] A. Borji, “Pros and cons of gan evaluation measures. arxiv 2018,” *arXiv preprint arXiv:1802.03446*,
- [33] P. Goyal, P. Dollár, R. Girshick, P. Noordhuis, L. Wesolowski, A. Kyrola, A. Tulloch, Y. Jia, K. He, “Accurate, large minibatch sgd: Training imagenet in 1 hour,” *arXiv preprint arXiv:1706.02677*, 2017.
- [34] A. Shrivastava. (2015). What’s the difference between gothic and romanesque architectures? [Online]. Available: <https://www.quora.com/Whats-the-difference-between-Gothic-and-Romanesque-architectures> (visited on 05/20/2020).
- [35] K. Richman-Abdou. (2018). Dazzling elements of ancient islamic architecture we still see today, [Online]. Available: <https://mymodernmet.com/islamic-architecture/> (visited on 05/20/2020).
- [36] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.
- [37] V. Nagarajan, C. Raffel, I. J. Goodfellow, “Theoretical insights into memorization in gans,” in *Neural Information Processing Systems Workshop*, 2018.
- [38] A. Krizhevsky, “Learning multiple layers of features from tiny images,” *University of Toronto*, May 2012.
- [39] F. Yu, A. Seff, Y. Zhang, S. Song, T. Funkhouser, J. Xiao, “Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop,” *arXiv preprint arXiv:1506.03365*, 2015.
- [40] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, *et al.*, “Imagenet large scale visual recognition challenge,” *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.

A

Derin Öğrenme Teknikleriyle Üretilen Görüntülerin Gerçekçiliğini Ölçme Anketi

Bu ankette, yapay bir nöron ağıyla üretilen, -daha önce hiç var olmamış- görüntülerin gerçekçi olup olmadığı test edilecektir.

Yapılan çalışmada yapay zekanın alt alanlarından birisi olan derin öğrenme yöntemleri kullanılarak İslam mimarisi ve Romanesk mimarisi örneklerini içeren çeşitli görüntüler üretilmiştir. Çalışma kapsamında, Çekişmeli-Üretici Yapay Nöron Ağları yaklaşık olarak 240.000 gerçek mimari yapı görüntüsü ile eğitilerek üretici bir model oluşturulmuştur. Eğitilen modelden İslam ve Romanesk mimarisine ait çeşitli örnekler üretilmiştir.

Aşağıda verilen görüntü örneklerini inceleyiniz. Sol sütunda yer alan görüntüler ağ tarafından üretilen örneklerdir. Sağ sütundaki görüntüler ise gerçek görüntülerdir ve üretilen örneğe en çok benzeyen gerçek görüntü örnekleri seçilmiştir.

ÜRETİLEN GÖRÜNTÜLERİN GERÇEK GÖRÜNTÜLERE BENZEMESİ ÖNEMLİ DEĞİLDİR. Gerçek görüntüler, eğitimde kullanılan örnekleri yalnızca temsil etmesi ve fikir vermesi amacıyla eklenmiştir. Üretilen örneklerin gerçekçi olup olmadığını değerlendiriniz. Gerçekçi (5) veya Gerçekçi değil (0) olarak işaretleyiniz. Üretilen görüntüler sol sütunda yer almaktadır.

Üretilen örneklerin gerçekçi olup olmadığını değerlendiriniz. Gerçekçi(5) veya Gerçekçi değil(0) olarak işaretleyiniz. Üretilen görüntüler sol sütunda yer almaktadır.



	0	1	2	3	4	5	
Gerçekçi Değil	☐	☐	☐	☐	☐	☐	Gerçekçi

Şekil A.1 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 1. soru

Soru *



0

1

2

3

4

5

Gerçekçi Değil

☐☐☐☐☐☐

Gerçekçi

Şekil A.2 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 2. soru

Soru *



	0	1	2	3	4	5	
Gerçekçi Değil	☐	☐	☐	☐	☐	☐	Gerçekçi

Şekil A.3 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 3. soru

Soru *



	0	1	2	3	4	5	
Gerçekçi Değil	☐	☐	☐	☐	☐	☐	Gerçekçi

Şekil A.4 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 4. soru

Soru *



	0	1	2	3	4	5	
Gerçekçi Değil	☐	☐	☐	☐	☐	☐	Gerçekçi

Şekil A.5 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 5. soru

Soru *



0 1 2 3 4 5

Gerçekçi Değil ☐ ☐ ☐ ☐ ☐ ☐ Gerçekçi

Şekil A.6 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 6. soru

Soru *



	0	1	2	3	4	5	
Gerçekçi Değil	☐	☐	☐	☐	☐	☐	Gerçekçi

Şekil A.7 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 7. soru

Soru



0

1

2

3

4

5

Gerçekçi Değil

☐☐☐☐☐☐

Gerçekçi

Şekil A.8 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 8. soru

Soru *



0

1

2

3

4

5

Gerçekçi Değil

☐☐☐☐☐☐

Gerçekçi

Şekil A.9 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 9. soru

Soru *

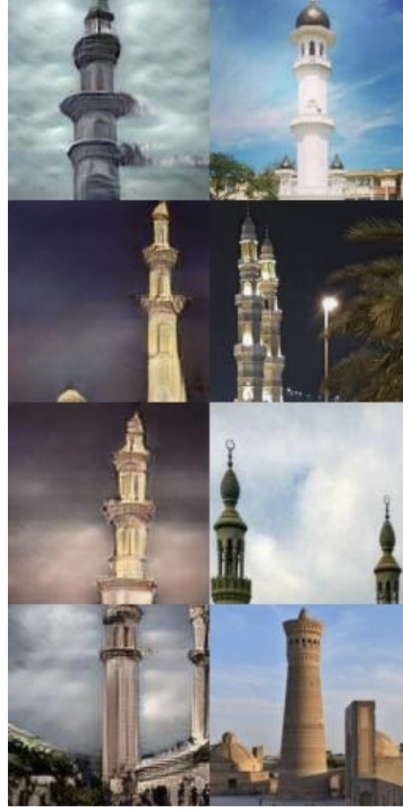


0 1 2 3 4 5

Gerçekçi Değil ☐ ☐ ☐ ☐ ☐ ☐ Gerçekçi

Şekil A.10 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 10. soru

Soru *



Gerçekçi Değil 0 1 2 3 4 5 Gerçekçi

Şekil A.11 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 11. soru

Soru *



	0	1	2	3	4	5	
Gerçekçi Değil	☐	☐	☐	☐	☐	☐	Gerçekçi

Şekil A.12 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 12. soru

Soru *



Gerçekçi Değil 0 1 2 3 4 5 Gerçekçi

Şekil A.13 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 13. soru

Soru *



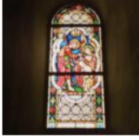
	0	1	2	3	4	5	
Gerçekçi Değil	☐	☐	☐	☐	☐	☐	Gerçekçi

Şekil A.14 Derin öğrenme teknikleriyle üretilen görüntülerin gerçekçiliğini ölçme anketi 14. soru

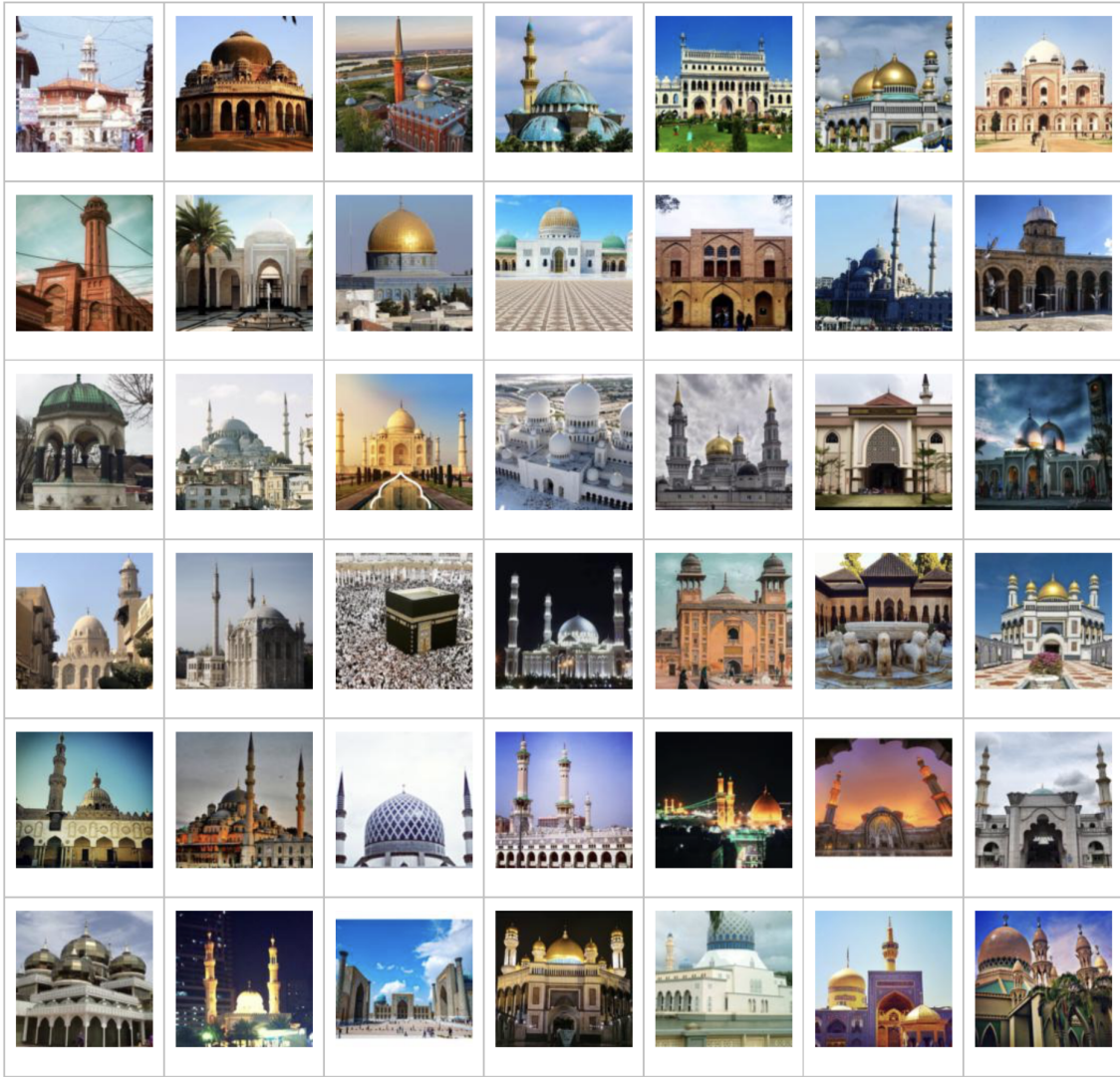
B Gerçek ve Üretilen Görüntü Veri Kümesi Örnekleri



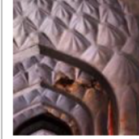
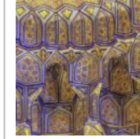
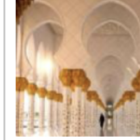
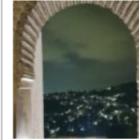
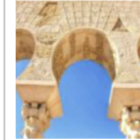
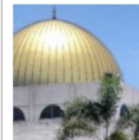
Şekil B.1 Veri kümesinde yer alan Romanesk mimarisine ait genel dış cephe örnekleri

Sunak						
Vitray						
Kule						
Tonoz						
Kapı Girişi						
Apsis						

Şekil B.2 Romanesk mimarisine ait veri kümesinde yer alan başlıca unsurlar



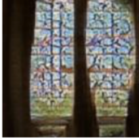
Şekil B.3 Veri kümesinde yer alan İslam mimarisine ait genel dış cephe örnekleri

Mukarnas						
Tonoz						
Kemer						
Minare						
Avlu Revakı						
Kapı Girişi						
Kubbe						

Şekil B.4 İslam mimarisine ait veri kümesinde yer alan başlıca unsurlar



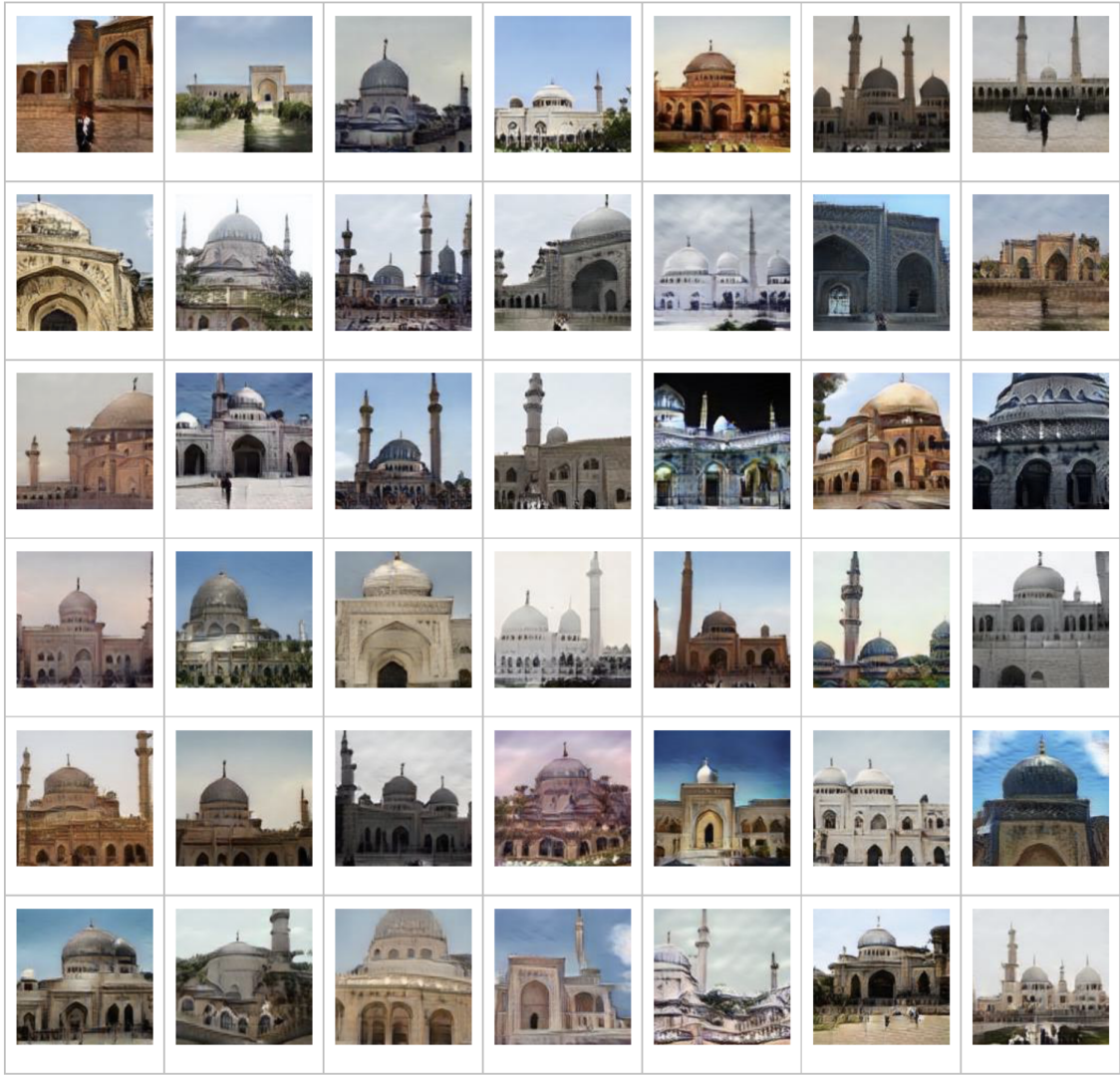
Şekil B.5 Model tarafından üretilmiş Romanesk mimari dış cephe örnekleri

Sunak						
Vitray						
Kule						
Tonoz						

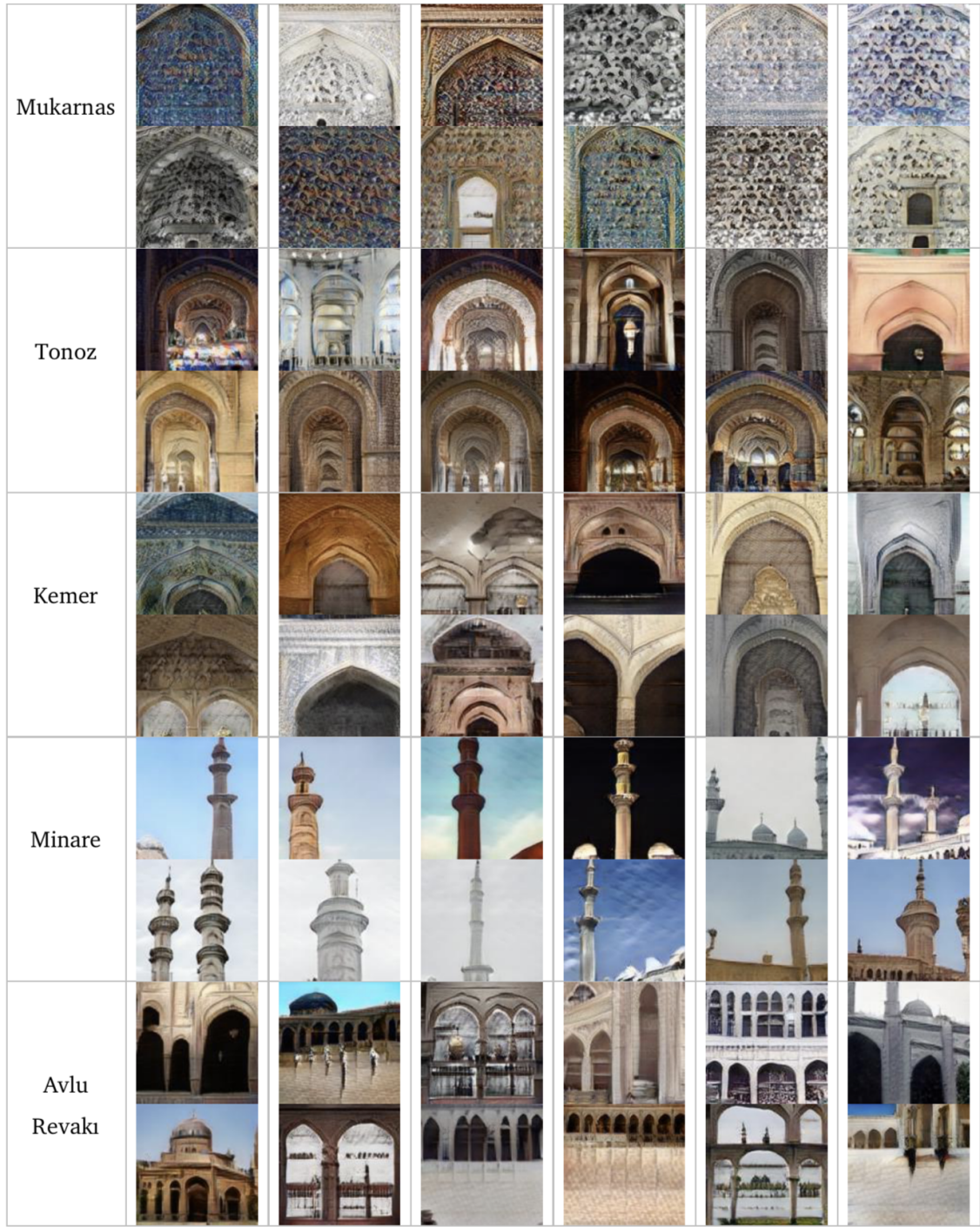
Şekil B.6 Romanesk mimarisine ait üretilmiş görüntülerde yer alan unsurlar

Kapı Girişi						
						
Apsis						
						

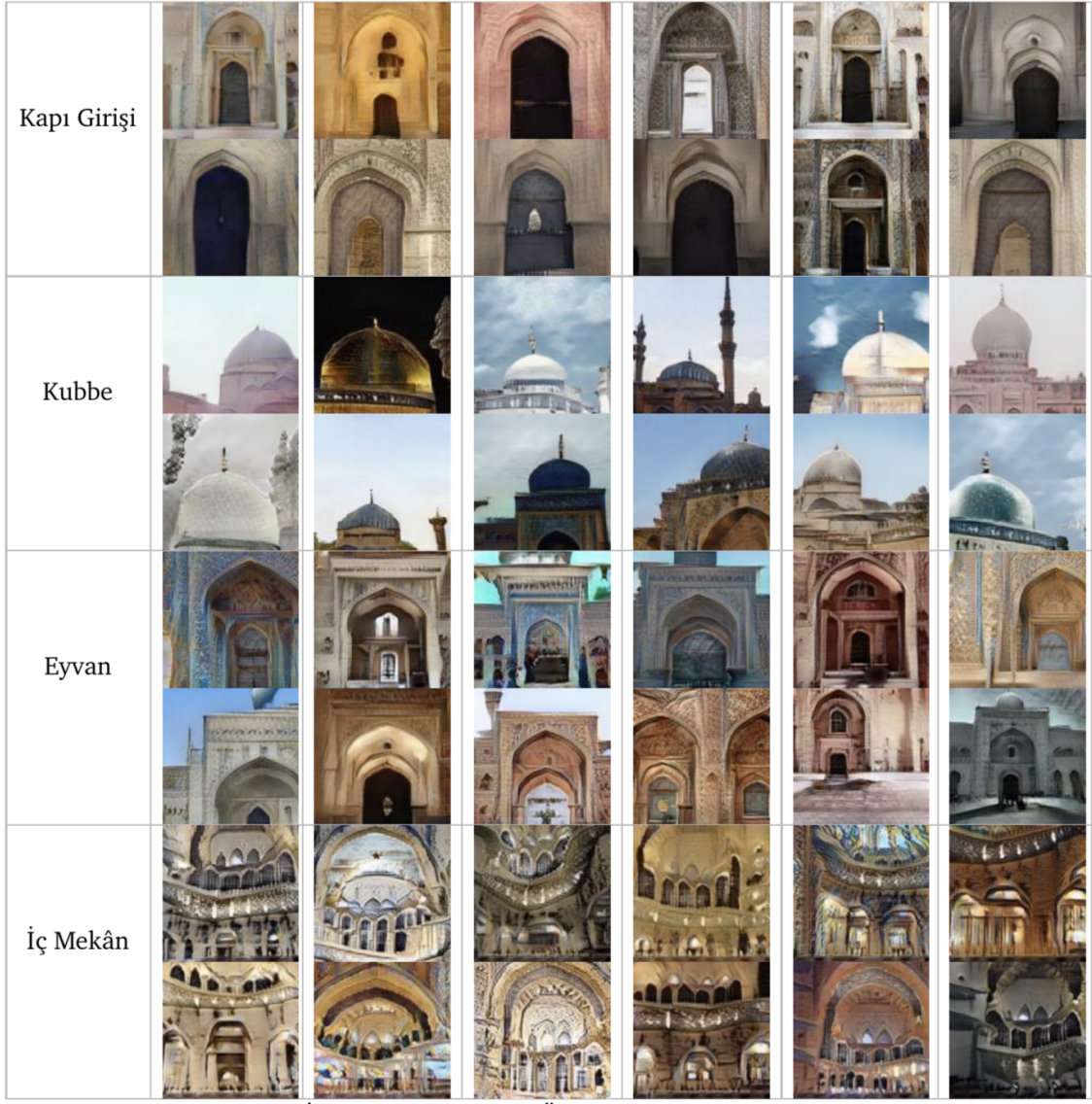
Şekil B.7 Romanesk mimarisine ait üretilmiş görüntülerde yer alan unsurlar



Şekil B.8 Model tarafından üretilmiş İslam mimarisi dış cephe örnekleri



Şekil B.9 İslam mimarisine ait üretilmiş görüntülerde yer alan unsurlar



Şekil B.10 İslam mimarisine ait üretilmiş görüntülerde yer alan unsurlar

TEZDEN ÜRETİLMİŞ YAYINLAR

İletişim Bilgileri: altuntasmukaddes@gmail.com.tr

Konferans Bildirisi

1. ISPEC 5th International Conference on EngineeringNatural Sciences, Full Texts Book, Volume-2, syf:550-568, ISBN: 978-605-7811-41-7

