

A Case Study on Evaluation of Architectural Plan Production with GAN

Can Uzun¹

0000-0002-4373-9732¹

¹ Istanbul Technical University, Graduate School of Science, Engineering, and Technology, Department of Informatics, Architectural Design Computing, Istanbul, Turkey

This study is a study on the evaluation of the methods by which the GAN algorithm is evaluated. Although GAN algorithm evaluation methods are accepted in the literature, it is still a question to be answered whether GAN evaluation methods can be used for evaluating the outputs of GAN training with a data set consisting of architectural plan schemes. Throughout this study, the architectural plan scheme outputs of GAN algorithm are evaluated. While making this evaluation, Frechet Inception Distance and Rapid Scene Classification were used, respectively, among the quantitative and qualitative evaluation methods accepted in the literature of the GAN algorithm. As a result of the evaluation, the suitability of these methods for the evaluation of autonomous architectural plan generation was discussed. At the end of the text, it is concluded that quantitative and qualitative GAN evaluation methods need new specialized methods to evaluate architectural plan scheme productions with GAN algorithm.

Received: 31.08.2020

Accepted: 04.10.2020

Corresponding Author:

uzunc@itu.edu.tr

Uzun, C. (2020). A Case Study on Evaluation of Architectural Plan Production with GAN. JCoDe: Journal of Computational Design, 1(3), 167-182.

Keywords: GAN Evaluation, Frechet Inception Distance, Rapid Scene Categorization, Autonomous Architectural Plan Scheme Generation

GAN ile Mimari Plan Üretimlerinin Değerlendirilmesi Üzerine Bir Durum Çalışması

Can Uzun¹

0000-0002-4373-9732¹

¹İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilişim Anabilim Dalı, Mimari Tasarımda Bilişim, İstanbul, Türkiye

Bu çalışma GAN algoritması değerlendirme yöntemlerin değerlendirilmesi üzerine bir çalışma niteliğindedir. GAN algoritması değerlendirme yöntemleri her ne kadar literatürde kabul görmüş olsa da mimari plan şemalarından oluşan bir veri seti eğitim çıktılarında da GAN verimliliğinin aynı değerlendirme yöntemleri ile kullanılıp kullanılmaması cevaplanması gereken bir soru halindedir. Bu çalışma boyunca GAN algoritmasının alt sınıfında bulunan DCGAN algoritması ile üretilmiş Palladyan plan şemalarının ve GAN algoritmasının verimliliği değerlendirilmiştir. Bu değerlendirme yapılırken GAN algoritmasının literatürde kabul görmüş nicel ve nitel değerlendirme yöntemlerinden sırasıyla Frechet Inception Distance ve Hızlı Sahne Sınıflandırması kullanılmıştır. Değerlendirme sonucunda bu yöntemlerin mimari plan üretimi için uygunluğu tartışılmıştır. Metnin sonunda nicel ve nitel GAN değerlendirme yöntemlerinin mimari plan şeması üretimlerini değerlendirmek üzere özelleşmiş yeni yöntemlere ihtiyacı olduğu sonucuna varılmıştır.

Teslim Tarihi: 31.08.2020

Kabul Tarihi: 04.10.2020

Sorumlu Yazar:

uzunc@itu.edu.tr

Uzun, C. (2020). GAN ile Mimari Plan Üretimlerinin Değerlendirilmesi Üzerine Bir Durum Çalışması. JCoDe: Journal of Computational Design, 1(3), 167-182.

Anahtar Kelimeler: GAN Değerlendirme, Frechet Inception Distance, Hızlı Sahne Sınıflandırması, Otonom Mimari Plan Şeması Üretimi.

1. GİRİŞ

Yapay zeka disiplinde yapay sinir ağlarına yönelik çalışmalar mimarlık disiplini ile etkileşime geçmeye başlamıştır. Bu etkileşimlerden biri GAN algoritmaları ile kurulmaktadır. 2014 yılında Ian Goodfellow tarafından önerilen GAN algoritmaları veriyi (görsel, işitsel) işleyerek veriye yüksek benzerlikte yeni üretimler gerçekleştirebilmektedir. (Goodfellow ve diğ., 2014). GAN algoritması, derin öğrenme algoritmalarının alt sınıfında yer alan üretken bir algoritmadır. GAN algoritması, ayrıştırıcı (discriminator) ve üretici (generator) iki fonksiyonun birbiri ile yarışması ile çalışmaktadır. Üretici fonksiyon ayrıştırıcı fonksiyonu ürettiği veri ile kandırabilirse üretimler algoritma tarafından gerçek olarak değerlendirilecek ve üretici fonksiyonun hata değeri düşmeye başlayacaktır. Bu şekilde üretici fonksiyon öğrendiği veriyi üretmeyi başarabilmektedir GAN algoritması veri üretim başarımını, görsel veriler üzerinde de oldukça verimli ve gerçekçi sonuçlar ile göstermiştir. GAN algoritmalarının görsel veriyi üretme gücü nedeniyle otonom mimari plan üretim çalışmalarında doğrudan yararlanılabilecek bir algoritma olmuştur (Huang & Zheng, 2018; Chaillou, 2019; Uzun ve diğ., 2020). GAN algoritmaları ile yapılan çalışmalar çoğunlukla piksel tabanlı imaj türünde mimari plan üretim çalışmalarını içermektedir (Huang & Zheng, 2018; Chaillou, 2019; Uzun ve diğ. 2020). Ancak otonom üretimlerin gerçekleştirilmesinin yanında diğer önemli nokta otonom üretimlerin değerlendirilmesidir. GAN ile üretilen mimari plan şemalarının değerlendirilmesi ile ilgili literatürde az sayıda çalışma bulunmaktadır (Chaillou, 2019; Uzun ve diğ. 2020). Bu durumun başlıca nedeni yapay sinir ağları ile mimarlık disiplini etkileşimi çalışmalarının oldukça yeni bir çalışma alanı olmasıdır. Güzelci ve diğ. (2019) yapay sinir ağları ile mimarlık disiplini arasındaki çalışmaların 2017 yılından itibaren artmaya başladığını göstermiştir. Aynı şekilde 2017 yılından itibaren GAN algoritmaları ile mimarlık disiplini arakesitinde çalışmalar görülmüştür, ancak literatürde hala gelişmekte olan bir alandır (Huang & Zheng, 2018; Chaillou, 2019; Uzun ve diğ. 2020).

GAN algoritma eğitimi başlı başına bir alanken, algoritmanın değerlendirilmesi GAN değerlendirme yöntemleri olarak farklı bir alan açmaktadır (Borji, 2019). GAN üretimlerinin değerlendirilmesi algoritmanın değerlendirilmesinden farklı bir ifadedir. GAN üretimleri çoğunlukla GAN değerlendirme yöntemleri ile değerlendirilmektedir (Borji, 2019). GAN değerlendirme yöntemleri literatüründen

anlaşılacağı üzere, eğer GAN algoritması değerlendirmesinde GAN algoritmasının performansı verimli ise GAN üretimleri de verimlidir şeklinde bir sonuca ulaşılmaktadır. Ancak mimari plan üretiminde GAN algoritmasının verimli çalışması her zaman doğru mimari plan şeması mı üretir sorusu yanıtlanmamış bir soru olarak bulunmaktadır. Mimari plan, GAN algoritma eğitimlerinde sıkça kullanılan fotoğraflarda piksel değerlerinden gelen salt görsel veri dışında, mekansal dizilim kararları verisini de içermektedir. GAN algoritması üretimleri değerlendirilmek üzere bir çok nitel ve nicel yöntemle sahiptir. Ancak bu değerlendirme yöntemleri iyi bir mimari plan üretiminin gerçekleştiğini gösterip göstermediği en önemli sorudur. Bu nedenle bu metnin amacı GAN algoritmasının nicel ve nitel değerlendirme yöntemlerinin, mimari üretim çıktılarının mimari anlamda değerlendirilmesinde yeterli olup olmadığının araştırılmasıdır. Bu bağlamda bu çalışma GAN ile mimari plan şeması üretimlerinin değerlendirilme yöntemlerinin bir değerlendirilmesi çalışmasıdır.

Bu değerlendirmenin gerçekleştirilmesi amacıyla GAN algoritmasının eğitiminin Palladyan plan şemaları ile gerçekleştirilmesi kararlaştırılmıştır. Bu seçimin nedeni Palladyan plan şemalarında sürekli tekrar eden plan kurgusu nedeniyle üretim çıktıları ile veri setinin karşılaştırılmasının daha kolay olabilmesidir. Böylece değerlendirme sürecinde karmaşıklığı az okunaklı plan şemaları karşılaştırılarak değerlendirilebilecektir.

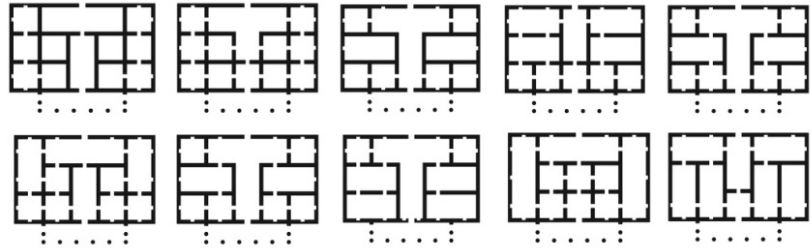
GAN eğitimi sürecinde ilk denenen veri seti gerçek Palladyan plan şemalarından oluşan bir veri seti olmuştur. Ancak gerçek Palladyan plan şemaları eğitim için yeterli sayıda veriyi oluşturamamıştır. Bunun yanında veri setinde bulunan gerçek Palladyan plan şemalarının çözünürlük kalitesi düşüktür. Bu nedenlerden dolayı GAN algoritması doğru bir genelleme yapamayarak çözünürlüğü ve okunurluğu düşük üretimler gerçekleştirmiştir. Bu üretimler GAN değerlendirme yöntemlerinde kullanılmak için yeterli kalitede olamamıştır. Bu nedenle gerçek Palladyan plan şemalarından oluşan veri seti ile gerçekleştirilen eğitim değerlendirilmemiştir. Veri setinin daha temiz ve yeterli büyüklükte olabilmesi için algoritma üretimi bir Palladyan plan şeması veri setinin üretilmesi kararlaştırılmıştır. Bu nedenle veri setinin üretiminde GRAPE-SGI (Grasl, T. (t.y.)) adında Palladyan plan şeması üreten bir biçim grameri yorumlayıcısı kullanılmıştır. GRAPE-SGI ile 320

adet Palladyan plan şeması üretilmiş ve GAN eğitimi bu veri seti ile gerçekleştirilmiştir.

Algoritma eğitim sürecinde GAN alt sınıfında bulunan DCGAN algoritması GRAPE-SGI ile türetilmiş 320 adet Palladyan plan şemalarından oluşan veri seti ile 15000 epok eğitilmiştir. Epok algoritma eğitimi sürecinde algoritmanın veri setinin tümünü kaç kere tekrar okuduğunun sayısal değeridir. GRAPE-SGI çevrimiçi

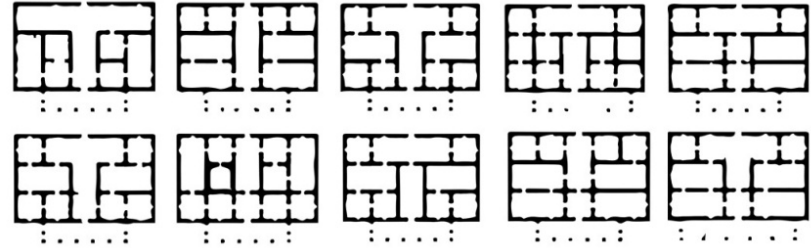
Şekil 1'de GRAPE-SGI ile oluşturulan Palladyan plan şeması veri setinden 10 plan şeması gösterilmiştir.

Şekil 1: GRAPE-SGI üretimi plan şemaları



15000 epok eğitilen DCGAN algoritmasının üretim çıktılarından 10 tanesi **Şekil 2'**de gösterilmiştir.

Şekil 2: DCGAN üretimi plan şemaları



İlk bakışta görseller birbirine benzese de, DCGAN'ın gerçekten verimli çalışıp çalışmadığı GAN değerlendirme yöntemleri ile test edilmelidir. Takip eden bölümlerde nicel ve nitel olarak eğitilen GAN algoritması değerlendirilmiş ve bu değerlendirme yöntemlerinin, üretilen mimari plan şemalarının değerlendirilmesindeki rolü tartışılmıştır.

2. GAN ALGORİTMASI DEĞERLENDİRME YÖNTEMLERİ

GAN algoritmaları ile mimari plan üretiminin otonom olabilmesi mümkün hale gelmiştir (Huang & Zheng, 2018; Chaillou, 2019; Uzun ve diğ., 2020). Üretimin gerçekleşmesi ile birlikte GAN algoritmasının üretim çıktılarının da değerlendirilmesi üretim verimliliğinin kararının

verilmesi için gerekli hale gelmektedir. Değerlendirme sonucunda hem GAN algoritması mimarisinin işlerliği ve verimliliği hem de üretim çıktılarının kalitesi değerlendirilmektedir. Üretimler değerlendirilmeden kullanılması halinde rastgele ve gerçek olmayan bir üretimin gerçekleşmiş olma ihtimali olacaktır. Algoritma üretim çıktılarının değerlendirilmesiyle algoritmanın hangi yönde üretimler yaptığı incelenebilir ve üretimlerdeki hatalı çıktılar temizlenebilir ya da hata payı büyük bir değerdeyse algoritma revize edilerek doğru yönelimde üretimlerin yapılması sağlanabilir. Bu nedenle algoritmanın üretim çıktılarının değerlendirilmesi, algoritma eğitim süreci kadar önemlidir.

GAN algoritmalarının üretim çıktılarının değerlendirilmesi yöntemleri ikiye ayrılmaktadır. Borji (2019), bu yöntemleri nicel ve nitel yöntemler olmak üzere ikiye ayırmaktadır. Nicel ve nitel yöntemlerin her biri kendi içinde avantaj ve dezavantajları da içermektedir.

Nicel yöntemlerden bazıları; Average Log-Likelihood, Coverage Metric, Inception Score, Modified Inception Score, Mode Score, Frechet Inception Distance, The Wasserstrain Critic, Geometry Score, Image Quality Measure, Precision Score ve F1 Score olarak sıralanabilir (Borji, 2019). Nicel değerlendirme yöntemlerinden veri seti ve üretim çıktısı arasındaki benzerliğin araştırıldığı yöntemlerde, logaritmik fonksiyonlar üzerinden tanımlanan olasılıksal yöntemler kullanılmaktadır. Böylece veri seti ve üretim çıktısı arasındaki benzerlik olasılıksal bir ifade ile tarif edilebilmektedir.

Nitel yöntemlerde GAN algoritması çıktıları insan gözlemci tarafından değerlendirilmektedir. Borji (2019), bu durumu insan yargılarının öznel olması nedeniyle dezavantajlı olduğunu söylemiştir ve nitel değerlendirmedeki üç problemi sıralamıştır. Borji'ye (2019) göre insan değerlendirmesi öznel yargılar nedeniyle yanlı olabilir bunun yanında yeteri kadar insan gözlemci bulma problemi ile karşılaşmaktadır. Bir diğer problem ise yine insanın öznel yargılarından kaynaklı insan gözleminin varyansının oldukça büyük olmasıdır. Başka bir deyişle herhangi bir veriden bir genelleme yaparak doğruya ulaşması ancak büyük bir veri seti ile karşılaşması sonucunda mümkün olabilir. Bu nedenle insan gözlemcinin değerlendirmesi sırasında karşılaştığı örneklem sayısı yüksek olmalıdır. Sonucu problem ise GAN algoritmasının mimarisi nedeniyle oluşan problemleri insan gözlemcinin tespit edebilecek algılarının bulunmamasıdır. Yüksek düzeyde işlem ve

hesaplama gerektiren bu sürecin insan tarafından yapılabilmesi mümkün olmayacaktır (Borji, 2019). Bu nedenle de GAN algoritmasının teknik problemleri nitel değerlendirme yöntemlerinde gözden kaçabilecek problemlerdir. Bu teknik problemler, aşırı uyum (overfitting), mod yığılımı (mode collapse) ve mod düşüşü (mod drop) şeklinde sıralanabilir. Aşırı uyum algoritmanın eğitim sırasında veri setini ezberlemesi ve ezberlediği veri dışında üretim yapmamış olmasıdır. Çözünürlük anlamında çok iyi olan bir GAN algoritma çıktısının veri setinden kopyalanmış, ezberlenmiş bir veri olduğunu insan gözlemci kaçırabilir. Mod yığılımı ise aynı görselin tekrar tekrar üretilmesi şeklindedir. Mod düşüşü ise üretimlerde hiç bir gerçek üretimin gözlemlenmemiş olması durumudur. Tüm bu problemler insan gözlemcinin gözünden kaçabilecek özellikleri barındırmaktadır (Borji, 2019). Bu nedenle nitel yöntemlerde değerlendirme sürecinde hatalar ile karşılaşılabilir. Nitel yöntemlerdeki bu dezavantaja rağmen algoritma üretim çıktılarının değerlendirilmesinde sıkça kullanılan bir yöntemdir. Hızlı sahne sınıflandırması (rapid scene categorization), tercih ile değerlendirme (preference judgement), en yakın komşu (nearest neighbor) nitel değerlendirme yöntemlerinde sınıflandırılmışlardır (Borji, 2019).

3. GAN ALGORİTMASI MİMARİ ÜRETİMLERİNİN DEĞERLENDİRMESİ

Bu bölümde GAN değerlendirme yöntemlerinden nicel ve nitel yöntemler kullanılarak elde edilen sonuçlar paylaşılacaktır. GAN ağlarını nicel değerlendirme yöntemlerinden Frechet Inception Distance (FID), nitel değerlendirme yöntemlerinden Hızlı Sahne Sınıflandırması (Rapid Scene Categorization) kullanılarak değerlendirme süreci ve sonuçları paylaşılacaktır.

3.1 FID ile Değerlendirme

Frechet Inception Distance GAN algoritması nicel değerlendirme yöntemleri arasında sınıflandırılmıştır. FID sentetik verinin dağılımı ile gerçek veri dağılımını karşılaştırarak sayısal bir sonuç vermektedir (Borji, 2019). FID yönteminin en önemli avantajı verinin gürültülü olması halinde de verimli çalışabiliyor olmasıdır. Ancak problemleri olan nokta FID yöntemi çalıştırılırken algoritma, karşılaştırılması yapılacak olan veriyi normal dağılımda (Gaussian / Normal distribution) kabul etmektedir (Borji, 2019).

Veri bilimcilerine göre gerçek dünyayı temsil eden her bir veri seti normal dağılıma (Gaussian / Normal distribution) sahiptir (URL1). Normal dağılım veriyi temsil eden değişkenlerin aritmetik ortalamasının sıfır olduğu dağılımdır. Görsel verilerden oluşan bir veri setinde özellikle siyah ve beyaz piksellerden oluşan mimari plan veri setinden veriyi temsil eden değişkenler (öznitelikler) siyah ve beyaz piksellerdir. Mimari planlardan oluşan bir veri setinin normal dağılıma sahip olabilmesi ise mimarlık disiplini içinden gelen bir diğer problem olacaktır. Mimari plan şemalarının tasarım eylemi nedeniyle çeşitlilik göstermesi normal dağılımı etkileme ihtimali olan bir özelliktir. Bunun yanında mimari planlardan oluşan büyük bir veri setinin henüz üretilmemiş olması da mimari plan verisini temsil edecek bir gerçek dünya modeli oluşturan veri setinin GAN algoritması eğitimlerinde kullanılamayacak olmasıdır. Tüm bu nedenlerden dolayı nicel olarak GAN algoritması eğitim çıktılarının değerlendirilmesi güvenilirliğinin kritik olduğu unutulmamalıdır. Ancak Borji (2019), normal dağılımla ilgili bu problemin her tür veri seti için değerlendirme verimini düşürebilecek bir özellik olduğunu belirtmiştir. Mimari plandaki özel durum ise normal dağılıma sahip bir veri setinin henüz üretilmemiş olmasıdır.

FID yönteminin çalışma prensibinde iki farklı veri kümesinin özniteliklerinin olasılıksal olarak dağılımlarının hesaplanması bulunmaktadır (Brownlee, 2019). FID veri setindeki gerçek veri ve DCGAN üretimindeki sentetik verinin her biri için öznitelik dağılımı hesaplamakta ve bu dağılımlar arasındaki mesafeyi bir skor olarak sonuçlandırmaktadır. Öznitelikleri veri içinde vektörlerle temsil edilmektedir. Bu vektörler veri içerisindeki değişkenlerdir. Gerçek görsel veri ve sentetik görsel verinin özniteliklerini temsil eden vektör değerleri arasındaki fark FID skorunu vermektedir. FID ismindeki mesafe (distance) bu iki veri arasındaki vektörel mesafe ile açıklanmaktadır. **Denklem 1**'de FID fonksiyonu gösterilmiştir (Brownlee, 2019).

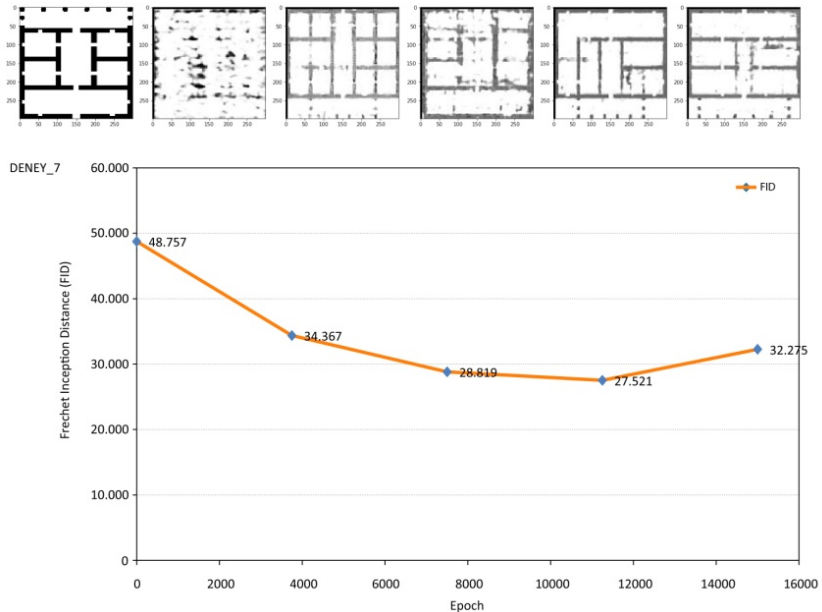
$$d^2 = ||\mu_1 - \mu_2||^2 + Tr(C_1 + C_2 - 2 * \sqrt{C_1 * C_2}) \quad (1)$$

Denklem 1'de mesafenin karesi d^2 ile ifade edilmektedir. C_1 ve C_2 ile kovaryans matrislerini göstermektedir. Veri seti ve DCGAN öznitelik dağılımları da μ_1 ve μ_2 ile temsil edilmiştir. Tüm bu fonksiyon ile sentetik ve gerçek veri arasındaki mesafenin sayısal değeri öznitelik

dağılımları kullanılarak elde edilebilir. GAN algoritmasının çalışmasının verimli olduğunu ifade etmek amacıyla her ne kadar FID değerinin olabildiğince küçük olması beklense de FID değerinin 0'a eşit çıkması algoritmanın veri seti ile birebir aynı üretim yaptığını göstermektedir. Başka bir deyişle yukarıda değindiğimiz algoritma eğitim problemlerinden aşırı uyum probleminin yaşandığı söylenebilir. FID skorunun 0 olması halinde GAN algoritmasının mimarisi ya da veri seti kaynaklı bir problem olduğu okunabilir. GAN algoritması değerlendirmesinde FID skoru kullanılırken beklenen değer ise her bir epok sonrasında FID skorunun sürekli düşen bir değer almasıdır. Böylece algoritmanın sürekli öğrenme yöneliminde ve kendini sürekli geliştirerek daha iyi temsiller ürettiği sonucuna varılabilir.

Şekil 3'te DCGAN algoritması kullanılarak üretilen Palladyan plan şemalarının GRAPE-SGI ile üretilen Palladyan plan şemalarından oluşan veri seti ile karşılaştırılması sonucu elde edilen FID değerlerini gösteren grafik verilmiştir.

Grafikte yatay eksen de epok düşey eksen de de FID değerleri okunabilir.



Şekil 3: DCGAN çıktısının Frechet Inception Score ile değerlendirilmesi.

Epok her bir eğitim basamağını ifade etmektedir. Tüm veri seti belli küme büyüklükleri ile belirli bir iterasyonda eğitilerek bir epok eğitimi tamamlar. Yani tüm veri setinin algoritmaya eğitilmesi 1 epok karşılığına

denk gelir. DCGAN' in eğitildiği bu otonom mimari plan üretimi deneyinde 15000 epok kadar eğitim sürdürülmüştür. Grafiğin yukarısında da DCGAN üretimi görsellerin epoklar sürecinde gelişimini göstermektedir. En baştaki görsel, veri setindeki gerçek veriyi ifade etmektedir. FID skoru grafiğine bakıldığında başlangıç epok ve sonuç epok arasında 11000. ve 15000. Epok arası dışında, sürekli bir azalma gözlenebilir. Ayrıca ilk epok ile son epok arasındaki değişim yaklaşık 16,5 birim kadardır. Bir diğer yandan epok artışı ile birlikte hem mekansal karşılıklarının olduğu hem de çözünürlük olarak daha kaliteli sonuçların son epokta elde edilebildiği gözlemlenebilir. 11000. epok sonrasındaki FID skorundaki artış, eğitimde veri setinin öznitelik dağılımındaki problemler nedeniyle olma ihtimali oldukça fazladır. Bir diğer yandan eğitimin 11000. epokta en verimli halini aldığı gözlenebilir. Bu da eğitimi 11000. epokta durdurmanın eğitim için yeterli olduğunu gösterir. Bu sonuçtan yola çıkarak FID skoru ile de eğitimin verimliliğinin her zaman daha fazla epok eğitiminin olması ile doğru bir orantısının olmadığı gösterilebilir. Ancak tüm bu değerlendirme sonrasındaki genel sonucumuz DCGAN'ın çoğunlukla doğru yönde eğitiminin devam ettiği ve veri setini doğru şekilde genelleyerek veri setine benzer üretimleri neredeyse iyi bir çözünürlük kalitesinde üretebildiği sonucuna ulaşabiliriz. Daha iyi bir özniteliksel dağılıma sahip veri seti ile DCGAN eğitiminde daha iyi bir başarımla alınabilse de DCGAN'ın mimari plan üretiminde verimli bir öğrenme süreci gerçekleştirdiğini bu verimlilik nedeniyle de dolaylı olarak mimari plan şemalarının uygun üretilmeye başlandığı söylenebilir.

Burada problemler olarak görülen, değerlendirme sonucu FID skoru üzerinden dolaylı olarak mimari plan üretimlerinin değerlendirilmiş olmasıdır. Sonuçta FID skoru GAN algoritmasının çalışma verimini matematiksel olarak doğru olduğunu çözümlenmiş olsa bile üretilen mimari planların mekansal organizasyonları ile ilgili bir bilgi vermemektedir. FID değeri mekansal dizilim ve mekanların fonksiyonları üzerine bir değerlendirme sonucu veremeyeceği için mimari plan organizasyonu değerlendirilmesinde işlevsiz bir değerlendirme yöntemi olacaktır.

Nicel yöntemlerle GAN algoritması değerlendirmesinin özellikle mimari plan görseli üretiminde kritik olarak ele alınması gereken bir yöntem olduğu düşünülmektedir. Bu nedenle bir sonraki bölümde nicel yöntem

değerlendirmesinin yanında nitel bir değerlendirme yöntemi de denenmiş ve çıktıları değerlendirilmiştir.

3.2 Hızlı Sahne Sınıflandırması Yöntemi ile Değerlendirme

GAN algoritması nitel değerlendirme yöntemleri arasında Hızlı Sahne Sınıflandırması (Rapid Scene Categorization) bulunmaktadır. Bu yöntem anket tabanlı bir yöntemdir. Hızlı sahne sınıflandırması insan gözlemcinin hızlı şekilde karar alarak sınıflandırma yapıyor olmasıdır. Hızlı sahne sınıflandırmasında gerçeklik değerlendirmesi yapılmaktadır. Gerçeklik değerlendirmesi, biri veri setinden diğeri ise GAN üretiminden iki görsel çıktının hangisinin gerçek olduğunun sorulması ve anket katılımcısı tarafından yanıtlanması ile gerçekleşir. Eğer katılımcı çoğunlukla GAN üretimlerini gerçek olarak sınıflandırırsa GAN algoritmasının verimli olduğu söylenebilir.

DCGAN ile üretilen Palladyan plan şeması üretimlerinin değerlendirilmesinde kullanılan nitel yöntem hızlı sahne sınıflandırması olmuştur. Hızlı sahne sınıflandırmasında, gerçek olan verinin bulunması hedeflenir. Ancak mimari plan şeması verisinde hangisinin gerçek hangisinin gerçek olmayan veri olduğunu belirleme durumu mimari planın bir özneliğini belirleyen bir seçim olmayacaktır. Bu nedenle seçilen soru gerçeklik özelliği üzerinden değil hangi plan organizasyonunun daha iyi olduğu sorusu olmuştur. Böylece anket katılımcıları GAN üretimlerinde plan organizasyonunu değerlendirmeye başlayacaklardır. Bu GAN algoritmasının verimliliğinden üretilen mimari plan şemalarının durumunun değerlendirilmesini sağlayacaktır.

Hızlı Sahne Sınıflandırması yapılırken **Şekil 1** ve **Şekil 2**'deki hem veri setinden hem de DCGAN üretimi 10'ar adet plan şeması görselinden oluşan on soruluk bir anket çalışması hazırlanmıştır. Anket SurveyMonkey üzerinden online gerçekleştirilmiştir. Her soruda bir DCGAN bir de GRAPE-SGI üretimi Palladyan plan şeması bulunmaktadır. Anket katılımcısı bu iki plan şeması üzerinden hangisinin daha iyi bir mekansal organizasyon sunduğunu seçmektedir. Katılımcı anket boyunca hangi verinin DCGAN, hangi verinin GRAPE-SGI üretimi olduğunu bilmemektedir.

Şekil 4 SurveyMonkey'de oluşturulan anket arayüzü ve bir örnek soruyu göstermektedir.

Plan Şeması Düzeni Değerlendirmesi / Plan Scheme Layout Evaluation

Plan Şeması Düzeni Değerlendirmesi / Plan Scheme Layout Evaluation

TR

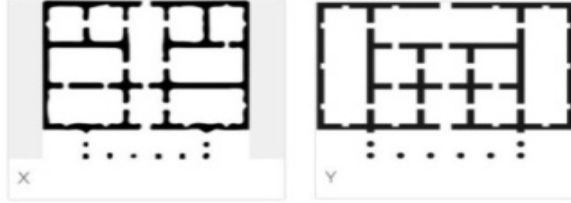
Bu ankette katılımcılardan aşağıdaki plan şemalarından hangilerinin daha doğru bir mimari plan düzenine sahip olduğunu işaretlemeleri beklenmektedir.

EN

In this survey, participants are expected to mark which of the following plan schemes have a more accurate architectural plan configuration.

Tamam

* 6. Hangi plan düzeni daha iyidir? / Which plan configuration is better?

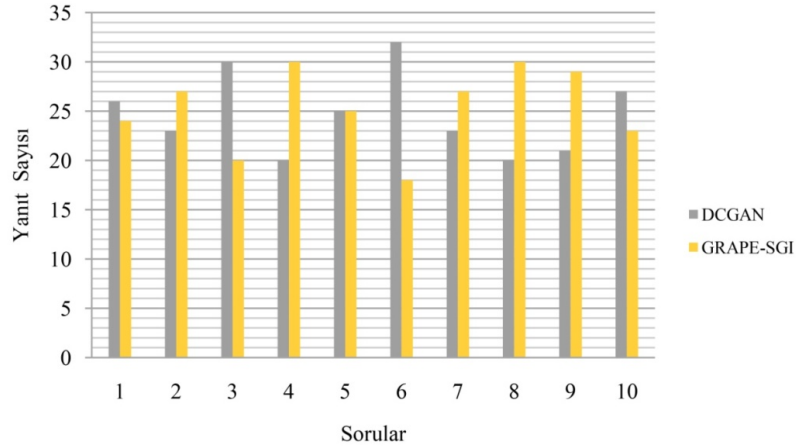


50 mimar katılımcı ile gerçekleştirilen bu ankette 10’ar plan şeması ile oluşturulan sorular üzerinden toplamda 500 farklı yanıt elde edilmiştir. Şekil 5’ te sorulara verilen yanıtların sayısal olarak dağılımı görülebilir.

Hızlı sahne değerlendirme sonucuna göre, Şekil 5’ teki grafikten de anlaşılacağı üzere sorulan 10 soru içinde 4 tanesinin DCGAN yanıtları GRAPE-SGI üretimi veri seti yanıtlarından daha fazla olmuştur. 1 soruda eşit sayıda yanıt alınmış ve diğer 5 soruda da GRAPE-SGI üretimleri daha yüksek sayıda yanıt almıştır. Ancak grafik üzerinde DCGAN ve GRAPE-SGI seçimleri her bir soru üzerinden incelendiğinde hiç bir soru için büyük bir fark olmadığı da okunabilir. En büyük sayısal farkın ise DCGAN üretimlerinin 14 yanıt fazla olduğu 6. sorudur.

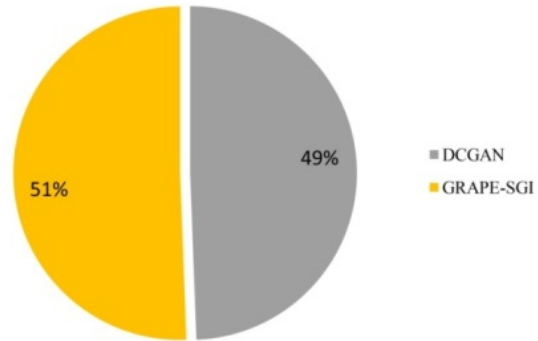
Şekil 4: Hızlı Sahne Sınıflandırması Anket Arayüzü

Şekil 5: Anket Yanıt Sayıları Grafiği



Şekil 6'da hızlı sahne sınıflandırması yanıtlarının yüzde dağılımı gösterilmiştir. GRAPE-SGI yanıtların %51'ini, DCGAN ise % 49'unu almıştır. Böylece katılımcılar plan düzenlemesi bakımından veri seti ve DCGAN üretim çıktılarının plan şeması organizasyonunu benzer düzeyde bulduğu söylenebilir.

Hızlı Sahne Sınıflandırması Yüzdeleri



Şekil 6: Anket Yanıt Yüzdeleri

Kalitatif yöntemlerden hızlı sahne sınıflandırmasına göre DCGAN üretimleri veri seti ile benzer bir mimari plan şeması organizasyonuna sahiptir. Ancak bu çalışmanın net bir sonucunun olabilmesi katılımcı sayısının ve soru sayısının daha fazla olması ile mümkün olabilir. Nitel yöntemlerin dezavantajlarında bildirilen katılımcı maaliyeti bu durumla ilgilidir. Bir diğer yandan katılımcıların seçimlerinde öznel yargılarının bulunması nedeniyle seçimlerinin değişiklik gösterebileceği de unutulmamalıdır. Hızlı sahne sınıflandırmasının diğer bir dezavantajı ise bu değerlendirme yönteminin verinin gerçekliği üzerine kurgulanıyor olmasıdır. Ancak mimari plan şemasında bir gerçeklik sorusunun doğru sonuca götürecek bir soru olmaması, başka bir deyişle gerçeklik kavramının mimari plan şeması özneteliği olmaması nedeniyle soru

değiştirilmiş ve hangi plan organizasyonunun daha iyi bir mekansal kurguya sahip olduğu sorulmuştur. Bu sorunun da kendi içinde veri setinin ve üretimlerin grid tabanlı üretimler olması nedeniyle problemler yanları olabilir. Ancak yine de bu soru DCGAN üretim verimliliğinden üretilen plan şemalarının karşılaştırılarak değerlendirilmesi bakımından nicel yöntemle göre daha faydalı sonuçlar verdiği söylenebilir.

4. SONUÇ

Bu metinde, GAN algoritmasının değerlendirme yöntemlerinin, GAN algoritması ile üretilen mimari plan şemalarının değerlendirilmesinde kullanımının verimliliği tartışılmıştır. GAN algoritması değerlendirme yöntemleri her ne kadar literatürde kabul görmüş olsa da mimari plan şemalarından oluşan bir veri seti ile eğitim çıktılarında da aynı değerlendirme yöntemlerinin kullanılıp kullanılmaması cevaplanması gereken bir sorudur.

Borji'nin (2019) sınıflandırmasını yaptığı GAN algoritması nitel ve nicel değerlendirmesi yöntemleri GAN algoritmasının çalışma verimliliğini ölçerek özellikle nitel yöntemlerde çoğunlukla dolaylı yoldan üretim çıktıları değerlendirilmektedir.

GAN algoritması değerlendirme yöntemi genel olarak üretim çıktılarının çözünürlük, netlik, gerçeklik ve hedef görsele benzerlik değerleri üzerinden bir analiz yöntemi listesi sunmaktadır. Ancak mimari plan üretimlerinde çözünürlük, netlik gibi özellikler mimari planın bir özelliğini göstermemekte, yalnızca mimari görselin temsil kalitesini açıklayabilmektedir.

Mimari planın *gerçeklik* niteliği bulunmamaktadır. Bu açıklama GAN algoritması tarafından üretilen bir kedi görseli için yapılabilmemesi mümkündür. Herhangi bir GAN algoritmasının doğru bir şekilde üretmediği kedi görseli gerçek bir kediyi temsil edemeyecektir. Ancak bir mimari plan için gerçek bir üretim ya da gerçek olmayan bir üretim denmesi mümkün olmayacaktır. Bu nedenle Hızlı Sahne Sınıflandırmasında hangi plan şeması gerçektir sorusu yerine hangi plan şeması organizasyonu daha iyi çözülmüştür sorusu katılımcılara sorulmuştur. Bu durum GAN algoritması nitel değerlendirme yöntemlerinin mimari plan üretimi değerlendirilmesinde yeni tür değerlendirme yöntemlerine ihtiyaç duyulduğunu göstermektedir.

GAN algoritması deęerlendirmesindeki dięer bir nitelik ise *benzerliktir*. Benzerlik GAN algoritmasının alıřma verimlilięi ile ilgili olup GAN algoritmasının veri setine birebir benzer bir grsel retmesi problemli olarak grlmektedir. Bu nedenle veri setindeki mimari planın birebir aynı řekilde retimi sorunlu olacaęı gibi, bir mimari planın veri setindeki mimari plan ile benzer olup olmadıęının belirlenmesi de dięer bir problemdir. retimi yapılan mimari planın benzerlięi rettięi stil zerinden mi, mekan kurgusu mu ya da sadece morfolojik bir benzerlik ile mi deęerlendirilmesi gerektięi nemli bir soru haline gelir. GAN aęının bir kedi grseli retimi yapması halinde o grselin veri setindeki kedi grsellerine ne kadar benzedięi kolayca anlařılabilir. Mimari bir plan iin benzerlięin hangi kavramlar zerinden aıklanacaęı bir deęerlendirme sorunu olarak grlmektedir.

Frechet Inception Distance yntemi ile nicel olarak GAN algoritması verimlilięi deęerlendirmesi veri seti ile ıktılar arasındaki benzerlik yakınlıęına bakmaktadır. Ancak FID tamamen matematiksel bir iřlem ile mimari plan řeması retimlerdeki zneteliklerin vektrel olarak mesafelerine bakmaktadır. Bir grselin znetelięi olarak piksellerin renk deęerlerinden oluřan bu vektrler FID skorunda nemli bir girdi olmasına raęmen, mimari plan řeması znetelikleri iinde deęildir. Bu nedenle aslında mimari plan řeması sadece piksel daęılımı benzerlięi zerinden deęerlendirilmiř, mekansal organizasyon zerinden bir deęerlendirme yapılmamıřtır. FID skoru, GAN algoritması ile mimari plan řeması retimi ıktılarının mekansal organizasyonunun deęerlendirilmesinde iřlevsiz kalmıřtır.

Hem nitel hem nicel deęerlendirmeler sonucunda GAN ile mimari plan řeması retimlerinin deęerlendirilmesinde, FID skorunun ve hızlı sahne sınıflandırmasının da kesin bir sonu alabilecek bir yntem olamadıęı gzlemlenmiřtir. Bu deęerlendirmenin doęru yapılması halinde algoritmanın mimarisi de doęru řekilde revize edilebilir ve otonom mimari plan retimi alıřmaları da daha verimli hale gelebilir. Sonu olarak GAN deęerlendirilmesi alıřmalarında, mimari plan řeması retimi deęerlendirilmesi zelinde yeni yntemlerin geliřtirilmesinin olduka nemli bir arařtırma alanı olduęu sylenebilir.

Referanslar

Borji, A. (2019). Pros and Cons of Gan Evaluation Measures. *Computer Vision and Image Understanding* 179, 41-65. <https://doi.org/10.1016/j.cviu.2018.10.009>.

Brownlee, J. (August 30, 2019). *How to Implement the Frechet Inception Distance (FID) for Evaluating GANs* [Web blog]. Eriřim tarihi: 20-11-2019, eriřim adresi: <https://machinelearningmastery.com/how-to-implement-the-frechet-inception-distance-fid-from-scratch/>

Chaillou, S. (2019). *AI+ Architecture: Towards a New Approach* [This Project]. Harvard University Graduate School of Design.

Grasl, T. (t.y.). GRAPE For Web - Shape Grammar Interpreter. Eriřim tarihi: 20-11-2019, eriřim adresi: <http://grape.swapzt.com/App/PalladianGrammar>.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative Adversarial Nets. *In Advances in Neural Information Processing Systems* (pp. 2672-2680). Arxiv: 1406.2661. Eriřim tarihi 20-11-2019, eriřim adresi: <https://arxiv.org/abs/1406.2661>.

Güzelci, O. Z., Alaçam, S., & Güzelci, H. (2019). Trend Topics and Changing Concepts of Computational Design in the Last 16 Years-A Content Analysis. *In Proceedings of the 37th Annual Conference of Education and research in Computer Aided Architectural Design in Europe (eCAADe) 1* (pp. 423-430). Porto, Portugal.

Huang, W., & Zheng, H. (2018). Architectural Drawings Recognition and Generation Through Machine Learning. *In Proceedings of the 38th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)*. Mexico City, Mexico.

Uzun, C., Çolakođlu, M. B., & İnceođlu, A. (2020). GAN as a Generative Architectural Plan Layout Tool: A Case Study for Training DCGAN With Palladian Plans, and Evaluation of DCGAN Outputs. *ITU A/Z 17 (2)*, 185-198. doi: 10.5505/itujfa.2020.54037.

Url-1: <https://en.wikipedia.org/wiki/Central_limit_theorem>, eriřim tarihi: 4-12-2019.