

# Kişi Görüntülerinin Nitelik Esaslı Üretilmesi

## Generating Person Images Based on Attributes

Mehmet Günel, Erkut Erdem, Aykut Erdem  
Bilgisayar Mühendisliği  
Hacettepe Üniversitesi  
Ankara, Türkiye  
{n14327319, erkut, aykut}@cs.hacettepe.edu.tr

**Özetçe** —Nitelik esaslı kişi görüntüsü üretimi; poz, cinsiyet, giyilen giysi tipi, çanta bulunup bulunmaması vb. belirli niteliklerde gerçekçi duran insan görüntülerinin üretilmesi problemidir ve bilgisayarlı görü alanında çok çeşitli uygulama alanları bulunmaktadır. Görüntü üretiminin gerçekleştirilmesi arka/ön plan, kısmi kapanma, kişinin duruşu, kamera bakış açısı ve mesafesi, nitelikler arası karmaşık ilişkiler, dengesiz ve düşük çözünürlüklü veri gibi nedenlerle oldukça zordur. İlgili çalışmalarda görece daha kolay veri kümelerinde daha kısıtlı nitelikler ile daha karmaşık modeller kurularak daha özel amaçlar için sentetik görüntüler üretilmiştir. Bu çalışmada, daha genel bir amaç düşünülerek şartlı çekişmeli üretici ağ tabanlı DCGAN-C modeli kurulmuş ve hem çoklu sınıf hem de çoklu etiket özelliği gösteren daha fazla nitelik ile sentetik kişi görüntüleri üretilmiştir. Sonuçta PA-100K veri kümesi üzerinde hem nicel hem de nitel deneyler yapılarak modelin performansı gösterilmiştir.

**Anahtar Kelimeler**—*Şartlı çekişmeli üretici ağlar, sentetik kişi görüntüleri*

**Abstract**—Attribute based person image generation is a problem which considers generating realistic person images using attributes like pose, gender, clothes, whether a bag is present or not etc. and it has wide variety of applications on computer vision. Realization of that generation process is quite difficult due to several reasons such as foreground/background, partial occlusion, stance of a person, camera angle and distance, complex relationships between attributes, unbalanced and poor quality data etc. Synthetic images are generated in related works which have worked on relatively easier datasets using less attributes with more complex models for more specific purposes. In this work, a more controversial goal was set up and a model named DCGAN-C is developed based on conditional generative adversarial networks and it can produce sythetic person images with both multi-class and multi-labels. Consequently, both quantitative and qualitative experiments were performed on the PA-100K dataset and the performance of the model was demonstrated.

**Keywords**—*Conditional generative adversarial networks, synthetic person images*

### I. GİRİŞ

Gerçekçi görünüme sahip nitelik tabanlı kişi görüntülerinin üretilmesi, bilgisayarlı görü açısından kişi yeniden saptama, görüntü düzenleme, kişisel niteliklerin tanınması gibi farklı konular için önemli bir araştırma konusudur. Ayrıca, görsel gözetim, moda, otonom araç, vb. uygulamalarda da kendine

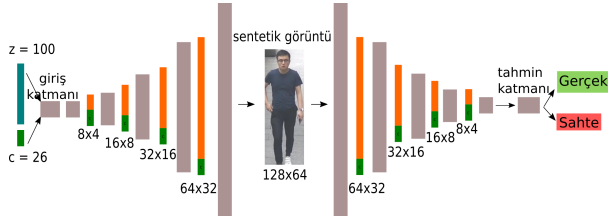
kullanım alanı bulmaktadır. Bahsi geçen bu uygulama alanları için kullanılan en başarılı yaklaşımların başında derin öğrenme tabanlı modeller geldiği düşünüldüğünde; bu modellerin eğitimde büyük miktarda eğitim verisine ihtiyaç duyulmaktadır. Eğitilecek bu modellere etiketli veri temin etmek ya da bilgisayar grafiği yardımıyla sentetik veriler sağlamak maliyetli ve zaman alıcıdır.

Bu çalışmada hem çevrimiçi giysi uygulamaları gibi konularda doğrudan, hem de kişi yeniden saptama, kişi nitelik tespiti gibi problemlerde kullanılmak üzere dolaylı olarak fayda sağlayabilecek nitelik esaslı sentetik kişi görüntülerinin üretilmesi problemi için bir araştırma gerçekleştirilmiş ve bu amaçla kullanılacak derin çekişmeli üretici ağlara dayanan yeni bir model geliştirilmiştir. Yaptığımız deneyler, geliştirilen modelin bell bir başarıyla, gerçekçi duran sentetik kişi görüntüleri üretebildiğini göstermektedir.

### II. İLGİLİ ÇALIŞMALAR

Derin öğrenme literatüründe son zamanlarda giderek önem kazanan bir konu, derin ağların gözetimsiz öğrenmede kullanım yollarının araştırılmasıdır. Bu bağlamda gerçekleştirilen çalışmaların önemli bir grubunu çekişmeli üretici ağlar olarak adlandırılan, yakın tarihte önerilmiş olan derin yapay sinir ağları oluşturmaktadır [3].

Sentetik kişi ve yüz görüntülerinin üretiminde bu model ve onun türevlerine dayanan bazı çalışmalar mevcuttur. Örneğin, DRAGAN [14] çalışmasında çeşitli nitelikler ile anime karakter yüzleri üretilmiştir. Bir başka çalışmada [15], modele çok daha fazla rehberlik edilerek ön/arka plan ve poz için ayrı ayrı maskeler kullanılmış ve görüntü üretimi için karmaşık bir model önerilerek nitelikler bunlarla sınırlandırılmıştır. AC-GAN [6] modelinde ise ayırt edici ağ ile aynı zamanda nitelik tahmini yapısı kurulmuş ve üretim kalitesini artırma amaçlanmıştır. AC-GAN çalışmasına paralel olarak DR-GAN [13] çalışmasında ise kimlik bilgisi korunmaya çalışılarak verilen yüz görüntüsünün çeşitli pozlardaki sentetik halleri üretilmiştir. SL-GAN [16] olarak önerilen çalışmada yüz resimlerinin çeşitli nitelikler ile üretilmesi/değiştirilmesi hedeflenmiştir. FM-GAN [17] yaş, cinsiyet ve etnik köken niteliklerini daha hassas kullanarak yüz resimlerine yönelmiştir. Konuya bir başka açıdan yaklaşan [18] çalışması ise sentetik veri üretimi ile kısıtlı verinin olduğu sınıflandırma problemi için GAN tabanlı öneride bulunmaktadır. Bu çalışmaya yakın olarak [19] çalışması da kişi yeniden saptama problemi için sentetik veriler üretmektedir. Ancak bu yöntemler çalışmamızda üzerine



Şekil 1: Çalışmamızda kullanılan DCGAN-C mimarisini

gittiğimiz nitelik esaslı üretimden farklı olup çalışmamız ile tam olarak örtüşmemektedirler.

Bu çalışmamızda ise bahsedilen diğer bütün çalışmalara göre kat kat fazla sayıda niteliğin kullanımı ile daha zorlu bir veri kümesi üzerinde daha basit bir model ile sentetik görüntü üretimi sağlanabilmiştir. Unutulmamalıdır ki artan nitelik ve sınıf sayısı ile GAN eğitimi oldukça zorlu bir hal alabilmektedir.

### III. YÖNTEM

#### A. Çekişmeli Üretici Yapay Sinir Ağları

Çekişmeli üretici sinir ağlar (Generative Adversarial Networks -GAN), yapay sinir ağlarının eğitim ve yapı açısından özel bir şekli olarak betimlenebilir. Bu ağ tipini ön plana çıkaran temel özellik, probleme uygun model ve iyi bir eğitim şartlarıyla, gerçeğe oldukça yakın görüntülerin üretimine olanak vermesidir.

Temel olarak GAN çatısı iki alt ağdan oluşmaktadır: 1. Üretici ağ ( $G$ ) ve 2. Ayırt edici ağ ( $D$ ). Üretici ağ  $G$ , belirli bir boyuta sahip rastgele bir  $z$  vektöründen gerçekçi sentetik görüntüler üretmeye çalışmaktadır. Ayırt edici ağ  $D$  ise üretici ağın oluşturduğu sentetik görüntüler ile gerçek görüntüleri birbirlerinden ayırt etmeyi amaçlayarak hangisinin gerçek hangisinin ise sahte olduğunu bulmaya çalışmaktadır.

Matematiksel açıdan Denklem 1’de görüleceği üzere,  $G$  ve  $D$  olarak nitelendirilen bu alt ağlar, aslında birbirleriyle çekişme halinde olup bir minimum-maksimum oyunu oynamaktadırlar.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log(D(x))] + \mathbb{E}_{z \sim p_Z(z)} [\log(1 - D(G(z)))] \quad (1)$$

Burada  $x$  gerçek bir girdi görüntüsünü ve  $p_{data}(x)$  ise gerçek görüntülerin dağılımını göstermekte iken,  $z$  bir rassal vektörü,  $p_Z(z)$  ise bu vektörlerin dağılımını temsil etmektedir.

Her ne kadar çekişmeli üretici ağları eğitmek pratikte oldukça zor olsa da, orjinal GAN [3] çalışmasında  $D$  ve  $G$  alt ağları uygun koşullarda birbirlerini izleyen değişimli güncellemeler ile eğitildiklerinde, üretici ağın  $p_{data}(x)$  gerçek görüntü dağılımına yakınsadığı teorik açıdan gösterilmiştir.

Belirtilen GAN yapısı DCGAN [1] çalışmasında evrimsel sinir ağlarının kullanımı ve eğitim aşamasında yapılan veri düzgeleme [12] gibi iyileştirmeler ile görüntü kalitesinde daha iyi, eğitimde ise daha dengeli hale getirilmiştir.

Bu çalışmada [1] çalışmasında belirtilen yapı baz alınarak DCGAN-C (DCGAN-Conditional) şartlı modeli çok sınıflı kişi niteliklerini dikkate alacak şekilde tanımlanmıştır.



Şekil 2: PA-100K veri kümesinden örnek görüntüler.

#### B. DCGAN-C

Çekişmeli üretici yapay sinir ağı parçası olan üretici ağ  $G$ , rassal bir  $z$  vektöründen yalnızca rastgele görüntüler üretebilmektedir. Görüntü üretimini istenilen şekilde yapabilmek amacıyla [4] ve [5] çalışmaları önerilmiştir. Bu çalışmalar ışığında, istenilen sınıfta görüntünün üretimi için,  $G(z)$  ve  $D(x)$  ağları,  $c$  sınıf bilgisini belirtmek üzere,  $G(z, c)$  ve  $D(x, c)$  şeklinde değiştirilmiş ve üretici  $G$  açısından  $p(x|c)$  dağılımı elde edilmeye çalışılmıştır. Ayrıca bu yolla eğitim daha da dengeli hale gelmiştir.

Bu çalışmada hedeflenen nitelik esaslı görüntü üretimi için sağlanan rassal  $z$  vektörüne,  $c$  çoklu etiket vektörü doğrudan eklenmiş ve bu eklenen  $c$  vektörü yalnızca  $G$  ağına ek girdi olarak sağlanmakla kalmamış, aynı zamanda  $D$  ve  $G$ 'nin tüm evrimsel sinir ağı katmanlarına eklenmiştir. Burada belirtilen  $c$  çoklu etiket vektörü, Veri Kümesi bölümünde belirtildiği gibi, *cinsiyet*, *yaş*, *poz*, *şapka*, *sırt çantası*, *vb.* nitelikleri barındırmaktadır. İlgili model görsel olarak Şekil 1 üzerinde belirtilmiştir.

Uygulanan bu yaklaşım ile özellikle veri kümesinde yeterli sayıda temsili bulunan nitelikler görüntü üretiminde etkin rol alabilmiştir. Bu yolla üretilen örnek sentetik görüntüler Şekil 3’te görülebilir.

### IV. DENEYLER

#### A. Veri Kümesi

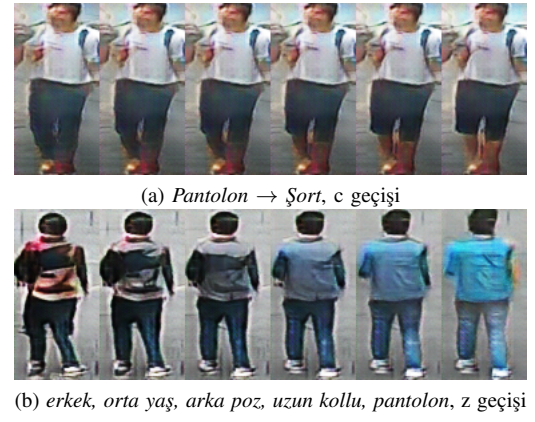
Gerçekleştirilen deneylerde PA-100K [2] veri kümesi kullanılmıştır. Bu veri kümesi, hem *cinsiyet*, *yaş*, *vb.* gibi global, hem de *sırt çantası*, *üst giyisi*, *vb.* gibi yerel nesne seviyesinde etiketlenmiş 80k eğitim, 10k doğrulama ve 10k test görüntüsü barındırmaktadır. Bu veri kümesinden rastgele seçilen örnek görüntülere Şekil 2’te yer verilmiştir. Bu şekilden de gözlemlenebileceği üzere veri kümesi belirtilen kişi nitelikleri cinsinden oldukça çeşitli kişi görüntüleri barındırmaktadır. Veri kümesinde resimlere ait toplamda 26 adet nitelik etiketlenmiş durumdadır. Bu nitelikler ve niteliklerin veri kümesindeki 100k görüntüde gözlemlenme sayıları Tablo I belirtilmiştir. Bu tablodan da görülebileceği üzere, çalışılan veri kümesi zorlu bir veri kümesi olup gerek resimlerdeki zorluklar gerekse dengesiz bir etiket dağılımı barındırmaktadır.

#### B. Görsel Sonuçlar

Geliştirilen DCGAN-C modelinin nitelik tabanlı sentetik kişi görüntüsü üretebilme başarısı görsel açıdan iki şekilde ele alınmıştır. İlk olarak nitelik bilgisi sabit tutularak, rastgele görüntüler üretilmiştir. Bu deneylerde  $z$  görüntü vektörü rastgele seçilirken, nitelik vektörü  $c$  sabit tutulmuştur. Bu sayede



Şekil 3: DCGAN-C ile üretilen örnek sentetik görüntüler ve nitelikleri.



Şekil 4: Rassal vektör  $z$  ve şart vektörü  $c$  üzerinde ilerleme.

Tablo I: PA-100K VERİ KÜMESİ NİTELİKLERİ VE SAYILARI

Nitelik Adı	Sayısı	Nitelik Adı	Sayısı
1 Kadın	54664	14 Kısa kollu	56913
2 60 yaş üzeri	1469	15 Uzun kollu	43087
3 18-60 yaş arası	92844	16 Üst atlamalı	5088
4 18 yaş altı	5687	17 Üst logo	14835
5 Ön poz	34707	18 Üst ekose	10917
6 Yan poz	30508	19 Üst bitişik	4219
7 Arka poz	34785	20 Alt çizgili	450
8 Şapka	4206	21 Alt desenli	1639
9 Gözlük	18662	22 Uzun palto	3365
10 El çantası	18115	23 Pantolon	71916
11 Omuz çantası	19301	24 Şort	16896
12 Sırt çantası	15926	25 Etek ve elbise	11155
13 Elde nesne tutma	958	26 Bot	595

aynı niteliklere sahip farklı kişi görüntüleri üretilmiştir. Bu deney ile elde edilen sonuçlar Şekil 3'te görülebilir.

Bunun ardından üretilen bir görüntünün yalnızca tek bir niteliği ile oynanarak bu değişiminin üretilen görüntü üzerindeki etkisi incelenmiştir. Bu deneyde gürültü vektörü  $z$  sabit tutulurken, nitelik vektörü  $c$  üzerinde istenilen değişiklikler yansıtılarak değiştirilmiştir ve bu ara noktalardan örneklemeler yapılmıştır. Şekil 4(a)'da bu yönde bir örnek verilmektedir. Benzer bir şekilde nitelik vektörü  $c$ 'nin sabit tutulurken gürültü vektörü  $z$  üzerinde değişikliklere gittiğimiz deneyler de gerçekleştirilmiştir. Şekil 4(b)'de yer verilen bu sonuçlardan gözlemlenebileceği üzere ara noktalardan elde edilen örneklemeler de aynı niteliklerde farklı kişi görüntülerine karşılık gelmektedir. Bu görsel örnekler, öğrenilen görüntü uzayının hem nitelik, hem de kişi özellikleri üzerinden öğreniminin başarımını göstermesi açısından önemlidir.

### C. Nicel Deneyler

Üretici modellerin başarılarının nicel olarak değerlendirilmesi konusunda Inception değeri [8], kullanıcı oylama sonuçları, MS-SSIM [7] gibi çeşitli yöntemler önerilmiştir. Ancak bu yöntemlerin güvenilirlikleri sorgulamaya açıktır [10] [7]. Bu nedenlerle günümüzde kabul görmüş yöntemler henüz bulunmamakla birlikte, bu alanda çalışmalar hala devam etmektedir.

Bu çalışmada modelin başarısının nicel olarak ölçülebilmesi için literatürden daha farklı bir yöntem izlenmiştir. Bu yöntemde öncelikle üretilen sentetik görüntüler ile nitelik tanımadaki kullanılan bir model eğitilmiş ve bunun ardından gerçek görüntüler ile eğitilen bu model üzerinde ince ayarlama (finetuning) yapılarak modelin başarısı ölçülmüştür. Bu deneyin sonucunda beklenen, üretilen sentetik görüntüler gerçek görüntülerin dağılımına uyduğu durum için, nitelik tanıma başarısında artma olması veya en azından başarıda negatif yönde herhangi bir değişimin olmamasıdır. Aksi durumda başarının düşmesi beklenmektedir.

Gerçekleştirdiğimiz deneylerde nesne tanımadaki kullanılan VGG16 [9] modeli esas alınmıştır. Bu modelin evrimsel filtreleri ImageNet yarışmasında eğitildiği biçimde kullanılmış, en sondaki nesne tanıma yönelik katman çoklu-etiket tanıma yönelik hale getirilmiş ve birbirini dışlayan nitelikler için (yaş, poz, cinsiyet, kısa/uzun kollu...) Softmax aktivasyon fonksiyonu kullanılmıştır. Ayrıca modelin son 2 tam katmanı, kendi problemimiz bu modelin geliştirildiği probleme göre daha basit olduğu için, 4096 boyutundan 2048 boyutuna düşürülmüştür.

Deney için eğitim kümesindeki niteliklerin dağılımı dikkate alınarak, DCGAN-C modeli ile rastgele 80k ve 400k adet sentetik görüntü üretilmiştir. Üretim için eğitim veri kümesindeki her bir niteliğin bu kümedeki oranı, sentetik veri üretiminde kullanılmak üzere bir seçim olasılığı olarak ele alınmış ve her bir nitelik için bağımsız seçim yapılarak nitelik kombinasyonları elde edilmiştir. Böylece nitelik vektörleri eğitim kümesi başta olmak üzere genel etiket dağılımını yakalamış, aynı zamanda eğitim kümesinde olmayan farklı kombinasyonlara da izin verir duruma gelmiştir.

Bu görüntüler ile ayrı ayrı eğitilen model, yalnızca eğitim kümesindeki gerçek görüntüler ile eğitime kıyasla test kümesinde küçük bir miktar artış göstermiştir. Ancak bu durumun, "sentetik olarak üretilen görüntüler test başarısını artırır" şeklinde yorumlanmaması gerekmektedir zira artış miktarı bu sonuca varabilmek için yeterli düzeyde gözükmemektedir. İlgili deney sonuçları Tablo II'de görülebilir. <sup>1 2</sup>

<sup>1</sup>Test aşamasında en iyi doğruluk başarı değerine sahip olan model kullanılmış, Adam algoritması  $\beta_1 = 0.5$  ve  $lr = 0.0002$  değerleri ile eğitilmiştir.

<sup>2</sup>Nitelik tahmininde üst ve alt giyisilerle ilgili nitelikler dikkate alınmamıştır.

Tablo II: NİTELİKLERİN TANINMASI SONUÇ TABLOSU

80k gerçek veri	80k gerçek +80k sentetik veri	80k gerçek +400k sentetik veri
0.870	0.874	<b>0.875</b>

Tablo III: TEST SONUCU NİTELİKLER VE DOĞRULUKLARI

80k Gerçek Veri		80k Gerçek + 400k Sentetik Veri	
Nitelik Adı	Doğruluk	Nitelik Adı	Doğruluk
Cinsiyet	0.782	Cinsiyet	<b>0.792</b>
Yaş	0.962	Yaş	<b>0.963</b>
Poz	0.723	Poz	<b>0.751</b>
Şapka	0.970	Şapka	<b>0.972</b>
Gözlük	0.791	Gözlük	<b>0.793</b>
El çantası	<b>0.859</b>	El çantası	0.857
Omuz çantası	<b>0.809</b>	Omuz çantası	0.798
Sırt çantası	0.839	Sırt çantası	<b>0.854</b>
Elde nesne tutma	<b>0.990</b>	Elde nesne tutma	<b>0.990</b>
Uzun/kısa kollu	0.790	Uzun/kısa kollu	<b>0.804</b>
Uzun palto	<b>0.956</b>	Uzun palto	<b>0.956</b>
Alt giysi	0.851	Alt giysi	<b>0.857</b>
Bot	<b>0.994</b>	Bot	<b>0.994</b>

Her ne kadar sonuç başarısında önemli bir değişim saptanmamış olsa da, Tablo III'te görülebileceği üzere, *cinsiyet*, *poz*, *sırt çantası* ve *uzun/kısa kollu* değerleri belirgin bir artış göstermiş, buna karşın diğer değerlerde dikkate değer bir değişim gözlemlenmemiştir. Bu durumun sebebi olarak, resmin daha genelini ilgilendiren ve resimde daha belirgin olan bu özelliklerin, geliştirilen DCGAN-C modeli ile daha iyi temsil edildiği düşünülmüştür.

#### D. Başarısız Deneyler

Yapılan başarılı deneylerin yanında, başarısızlıkla sonuçlanan deneylere de bildirimizde kısaca yer vermeyi diğer araştırmacılara ışık tutması açısından uygun bulduk. DCGAN-C modeline ek olarak, son katmanda gerçek/sahte tahmininin dışında, nitelik tahmini ile ilgili katmanlar da eklenerek DCGAN-A (DCGAN-auxiliary) modeli kurulmuştur. Bu modelin eğitimi *SigmoidCrossEntropy* ile çeşitli niteliklerle denenmiş ancak mütemediyen "mod çökmesi" problemi ile karşılaşmıştır. Ayrıca modelin *SigmoidCrossEntropy* ile eğitimi dışında, her bir nitelik için ayrı ayrı kayıp fonksiyonları tanımlayarak *SoftmaxCrossEntropy* fonksiyonu ile eğitimi gerçekleştirme yoluna da gidilmiştir. Deneylerimizde gradient değerlerinde patlama görülmüş, bunun önüne geçebilmek için yığılma veri düzgeleme [12], tek tek sıralı olarak kayıp fonksiyonlarını güncelleme, Wasserstein uzaklığı kullanma [11], *c* etiketlerine gürültü ekleme metodları uygulanmış ancak yine başarısız sonuçlar elde edilmiştir. Ek olarak DCGAN-C ile sentetik görüntüler oluşturulurken eğitim kümesindeki niteliklerin dağılımı dikkate alınmadan tamamen rastgele oluşturulan görüntüler gerçeklikten uzak görünmüşlerdir.

#### V. SONUÇ VE GELECEK ÇALIŞMALAR

Bu çalışma ile çekişmeli üretici yapay sinir ağları tabanlı bir model kurularak, kişi görüntülerinin çeşitli nitelikler ile grafik tabanlı üretim yöntemlerine göre daha az eforlu olacak biçimde veri bazı üretilebileceği gösterilmiştir. Bu anlamda belirli nitelikler ile belirli bir kalitede veri kümeleri elde edilebilmiş ve bu kümelerin nitelik tanıma eğitiminde başka modeller ile birlikte kullanılabilirliği de vurgulanmıştır.

Gelecekteki çalışmalarda bu çalışmanın devamı olarak; görüntülerdeki kalitenin artırılması, üretimde kullanılan niteliklerin olabildiğince birbirlerinden ayrık hale getirilmesi, kişi kimlik bilgisinin de dikkate alınması ve ince detayların daha belirgin hale getirilmesi ve başka modellerin eğitiminde etkin biçimde kullanılabilirliği ve başarıyı daha ileri götürebilecek üretim tekniklerinin araştırılması gibi hedeflerimiz bulunmaktadır.

#### TEŞEKKÜR

Bu çalışma kısmen TÜBİTAK 217E029 nolu proje tarafından desteklenmiştir. Ayrıca NVIDIA'ya sağladıkları GPU kartı için teşekkür ederiz.

#### KAYNAKÇA

- [1] A. Radford, L. Metz, S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks", arXiv:1511.06434, 2015
- [2] Xihui Liu, Haiyu Zhao, Maoqing Tian, Lu Sheng, Jing Shao, Shuai Yi, Junjie Yan, Xiaogang Wang, "HydraPlus-Net: Attentive Deep Features for Pedestrian Analysis", ICCV, 2017
- [3] Ian J. Goodfellow, Jean Pouget Abadie, Mehdi Mirza, Bing Xu, David Warde Farley, Sherjil Ozair, Aaron C. Courville, Yoshua Bengio, "Generative Adversarial Networks", NIPS, 2014
- [4] Mehdi Mirza, Simon Osindero, "Conditional Generative Adversarial Nets", arXiv:1411.1784, 2014
- [5] Hanock Kwak, Byoung Tak Zhang, "Ways of Conditioning Generative Adversarial Networks", arXiv:1611.01455, 2016
- [6] Augustus Odena, Christopher Olah, Jonathon Shlens, "Conditional Image Synthesis With Auxiliary Classifier GANs", ICLR, 2017
- [7] Kede Ma, Qingbo Wu, Zhou Wang, Zhengfang Duanmu, Hongwei Yong, Hongliang Li, Lei Zhang, "Group MAD Competition? A New Methodology to Compare Objective Image Quality Models", CVPR, 2016
- [8] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen, "Improved Techniques for Training GANs", NIPS, 2016
- [9] Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", ILSVRC, 2014
- [10] Shane Barratt, Rishi Sharma, "A Note on the Inception Score", arXiv:1801.01973, 2018
- [11] Martin Arjovsky, Soumith Chintala, Léon Bottou, "Wasserstein GAN", arXiv:1701.07875, 2017
- [12] Sergey Ioffe, Christian Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift", ICML, 2015
- [13] Luan Tran, Xi Yin, Xiaoming Liu, "Disentangled Representation Learning GAN for Pose-Invariant Face Recognition", CVPR, 2017
- [14] Yanghua Jin, Jiakai Zhang, Minjun Li, Yingtao Tian, Huachun Zhu, Zhihao Fang, "Towards the Automatic Anime Characters Creation with Generative Adversarial Networks", arXiv:1708.05509, 2017
- [15] Liqian Ma, Qianru Sun, Stamatios Georgioulis, Luc Van Gool, Bernt Schiele, Mario Fritz, "Disentangled Person Image Generation", arXiv:1712.02621, 2017
- [16] Weidong Yin, Yanwei Fu, Leonid Sigal, Xiangyang Xue, "Semi-Latent GAN: Learning to generate and modify facial images from attributes", arXiv:1704.02166, 2017
- [17] Lipeng Wan, Jun Wan, Yi Jin1, Zichang Tan, Stan Z. Li, "Fine-grained Multi-attribute Adversarial Learning for Face Generation of Age, Gender and Ethnicity", ICB, 2018
- [18] Antreas Antoniou, Amos Storkey, Harrison Edwards, "Data Augmentation Generative Adversarial Networks", ICLR 2018 Workshop, 2018
- [19] Zhedong Zheng, Liang Zheng, Yi Yang, "Unlabeled Samples Generated by GAN Improve the Person Re-identification Baseline in vitro", ICCV, 2017