



Kontrastlı Öğrenme Tabanlı Çekişmeli Üretken Ağlar İle Papsmear Görüntü Bölütleme

Papsmear Image Segmentation with Contrastive Learning Based Generative Adversarial Networks

Sara Altun Güven  Muhammed Fatih Talu 

¹İnönü Üniversitesi, Bilgisayar Mühendisliği Bölümü, Malatya / Türkiye
(kazimhanbay@gmail.com)

Received: Dec.19, 2021

Accepted: Jan.29, 2022

Published: Jun.06, 2022

Özetçe— PapSmear görsellerinin otomatik olarak rahim ağzı kanser varlığının tespit edilmesi aktif bir çalışma alanıdır. PapSmear görüntülerinde nesnelerin dağılımı sürekli yer değiştirmektedir. Bu çalışmada, Çekişmeli Üretken Ağlar (ÇÜA) ve karşılaştırmalı öğrenme tekniklerinden parça tabanlı yöntemler kullanılarak PapSmear görüntü bölütlemesi yapılmıştır. Kıyaslanan yöntemler CycleGAN, CUT, FastCUT, DCLGAN ve SimDCL yöntemidir. Tüm yöntemler eşlenmemiş görüntüler üzerinde çalışmaktadır. Bu yöntemler bir birlerini temel alarak geliştirilmişlerdir. DCLGAN ve SimDCL yöntemi CUT ve CycleGAN yönteminin birleşimidir. Bu yöntemlerde maliyet fonksiyonları, ağ sayıları değişkenlik göstermektedir. Bu çalışmada yöntemler ayrıntılı bir şekilde incelenmiştir. Yöntemlerin birbirine benzerlik ve farklılıkları gözlemlenmiştir. Bölütleme yapıldıktan sonra hem görsel hem de ölçüm metrikleri kullanılarak bulunan sonuçlara yer verilmiştir. Ölçüm metriği olarak FID, KID, PSNR ve LPIPS yöntemleri kullanılmıştır. Yapılan deneysel çalışmalar, DCLGAN ve SimDCL yönteminin PapSmear bölütletmede kıyaslanan yöntemler arasında daha iyi oldukları olduğu gözlemlenmiştir. CycleGAN yönteminin ise diğer yöntemlerden daha başarısız olduğu gözlemlenmiştir.

Anahtar Kelimeler: Görüntü bölütleme, derin öğrenme, evrimsel sinir ağı, çekişmeli üretken ağlar, karşılaştırmalı öğrenme

Abstract— Automatically detecting the presence of cervical cancer in PapSmear images is an active field of study. The distribution of objects in PapSmear images is constantly changing. In this study, PapSmear image segmentation was performed by using patch-based methods from Generative Adversarial Networks (GAN) and contrastive learning techniques. The methods compared are CycleGAN, CUT, FastCUT, DCLGAN and SimDCL methods. All methods work on unpaired images. These methods were developed on the basis of each other. DCLGAN and SimDCL method is a combination of CUT and CycleGAN methods. In these methods, cost functions and network numbers vary. In this study, the methods were examined in detail. Similarities and differences between the methods were observed. After segmentation, the results obtained using both visual and measurement metrics are included. FID, KID, PSNR and LPIPS methods were used as measurement metrics. Experimental studies have shown that DCLGAN and SimDCL method are better among the compared methods in PapSmear segmentation. It has been observed that the CycleGAN method is more unsuccessful than other methods.

Keywords: Image segmentation, deep learning, convolutional neural network, adversarial generative networks, contrastive learning

1.Giriş

Görüntü sentezleme bilgisayarla alakalı görsel alanlarda geniş kapsamda kullanılmaktadır. Sentezleme işlemi yapılırken bir alandaki görselin diğer alandaki görsel ile eşlemeyi öğrenmelidir.

Derin öğrenme algoritmalarından olan Çekişmeli Üretken Ağlar (ÇÜA), görüntü sentezlemede 2014 yılında Ian Goodfellow tarafından önerilmiştir (Goodfellow vd., 2014). ÇÜA'lar görüntü sentezleme yaparken denetimli ve denetimsiz olarak dönüşüm gerçekleştirirler. Denetimli dönüşüm yapılırken iki farklı alanda görüntü çifti bulunması gerekir. Olasılık dağılımını öğrenmek için her giriş görüntüsü ile buna karşılık gelen diğer alandaki çıkış görüntüsüne dönüşüm yapar. Pix2Pix (Isola vd., 2017), Pix2PixHD (Wang vd., 2018), StarGAN (Choi vd., 2018), SB-GAN (Azadi vd., 2019), BicycleGAN (Zhu vd., 2017) ve CEGAN (Xiong vd., 2019) yöntemleri denetimli dönüşüme örnektir. Denetimsiz dönüşüm ise alanlar arasında eşleme yaparken eşleştirilmiş görüntüye ihtiyaç duymamaktadır. Denetimsiz dönüşüme CycleGAN (Zhu vd., 2017), DiscoGAN (Kim vd., 2017) ve AttentionGAN (Mejjati vd., 2018), UNIT (Liu vd., 2017) ve BranchGAN (Zhou vd., 2019), MUNIT (Huang vd., 2018) ve DosGAN (Lin vd., 2019) yöntemleri örnek verilebilir.

CUT (Contrastive Unpaired Translation) yöntemi Park ve diğerleri tarafından 2020 yılında parçacık tabanlı öğrenme olarak önerildi. Bu yöntem, denetimsiz görüntüden görüntüye dönüşüm için karşılaştırmalı öğrenmeyi kullanmayı daha faydalı hale getirerek, yeni karşılaştırmalı öğrenme tekniğidir. DCLGAN (Dual Contrastive Learning for Unsupervised Image-to-Image Translation) yöntemi, CUT yöntemindeki mod çökmesini verimli şekilde gidermektedir. Bu yöntemde hem denetimsiz hem de denetimli yöntemler arasındaki boşluğun verimli bir şekilde kapatılabileceği gösterilmiştir. DCLGAN, denetimsiz dönüşüm yapan CycleGAN yöntemi ile CUT yönteminin birleştirilmesiyle oluşturulan yeni bir yaklaşımdır.

Medikal görüntü bölütlemesi hala büyük bir araştırma alanıdır. Derin sinir ağları, konvolüsyonel sinir ağları ve ÇÜA'lar görüntü bölütlemesi yapılırken örnek olarak kullanılan yapılarıdır (Yurt vd., 2021; Yao vd., 2021).

Bu çalışmada, görüntü bölütlenirken serviks kanserinin teşhisinde kullanılan PapSmear görüntüleri kullanılmıştır. Serviks kanseri bayanlar arasında dünyada yer alan ikinci ölümcül kanserdir (Shakuntala, 2016). Rahim ağzı kanseri teşhis ve tanısı yapılırken PapSmear hücre görüntülerini sınıflandırmak ve tespit etmek için yeni yöntemler önermişlerdir (Mustafa vd., 2020; Fekri-Ershad, 2019 ; Gautham vd., 2018). Mustafa vd., 2020 yılında çekirdek sınıflandırmasına ve bölütlenmesine dayanan servikal algılamanın kapsamlı şekilde incelemesini yapmıştır (Mustafa vd., 2020). Fekri Ershad, 2019 yılında PapSmear sınıflandırmasındaki eksiklikleri gidermek için görüntü döndürmede ve gri değer aralığı değişimlerine karşı hemen çökmeyen dirençli bir yöntem önermiştir (Fekri Ershad, 2019). Gautham vd., 2018 yılında tek hücreli çekirdek görüntülerinin sınıflandırılması için karar ağacına dayanan ve bu görüntülerin bölütlenmesinde ise CNN kullanan parça tabanlı bir yaklaşım önermişlerdir.

Bu çalışmada, yukarıda bahsi geçen PapSmear görüntü bölütleme için yeni ÇÜA mimarileri ve karşılaştırmalı öğrenme yöntemleri kıyaslanmıştır. ÇÜA mimarilerinden CycleGAN (Zhu vd., 2017), DCLGAN (Han vd., 2021) ve SimDCL (Han vd., 2021) yöntemleri; Karşılaştırmalı öğrenme yöntemlerinden ise görüntülerin parçalarını karşılaştırarak sonuca giden parça tabanlı karşılaştırma (CUT) yöntemi (Park vd., 2020) ile CUT yönteminin daha hızlı çalışmasını sağlayan FastCUT yöntemi ile çalışılmıştır. Yapılan metriksel karşılaştırmalar sonucu en son çıkan DCLGAN yöntemi bölütleme yapılırken diğer yöntemlere göre daha üstün sonuçlar elde ettiği gözlemlenmiştir.

2. Yöntemler

2.1. CycleGAN

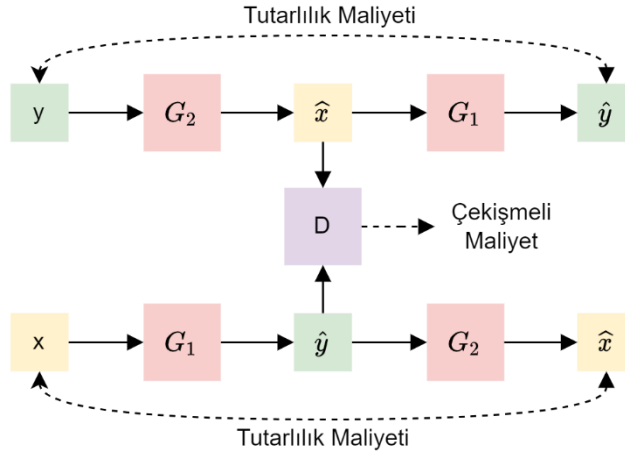
CycleGAN yöntemi çift taraflı ve birbirinden bağımsız iki ÇÜA mimarisi ile birlikte çalışmaktadır. X alanındaki bir giriş görüntüsünü Y alanındaki bir çıkış görüntüsüne dönüştürürken $G: X \rightarrow Y$ ve $F: Y \rightarrow X$ şeklinde iki ÇÜA mimari eğitimi gerçekleştirilir (Zhu vd., 2017). Buna göre maliyet fonksiyonu aşağıdaki gibi yazılır:

$$\begin{aligned} G^*, F^* &= \underset{G, F}{\operatorname{argmin}} \underset{D_X, D_Y}{\operatorname{max}} \mathcal{L}(G, F, D_X, D_Y) \\ &= \mathcal{L}_{GAN}^G(G, D_Y, X, Y) \\ &\quad + \mathcal{L}_{GAN}^F(F, D_X, Y, X) \\ &\quad + \lambda \mathcal{L}_{cyc}(G, F) \end{aligned} \quad (1)$$

\mathcal{L}_{cyc} , döngüsel tutarlılık maliyetidir. Döngüsel maliyet hesaplanırken $X \rightarrow \hat{Y} \rightarrow \hat{X}$ ve $Y \rightarrow \hat{X} \rightarrow \hat{Y}$ dönüşümleri yapılır. $X - \hat{X}$ ve $Y - \hat{Y}$ değerleri en küçük seviyeye çekilerek en iyi görüntü dönüşümünün sağlanması istenir. Buna göre döngüsel maliyet aşağıdaki gibi hesaplanır:

$$\mathcal{L}_{cyc} = |G(Y) - Y| + |F(X) - X| \quad (2)$$

CycleGAN yöntemi toplu normalleştirme yerine örnek (instance) normalleştirme kullanmaktadır. Üretici ağı yapılarında ise ResNET tabanlı yeni bir mimari kullanmaktadır. Şekil 1'de CycleGAN mimarisi görülmektedir. Bu mimaride iki üretici ağ ve bir ayırıştırıcı olduğu gösterilmektedir.



Şekil 1. CycleGAN mimarisi

2.2. Parça Tabanlı Karşılaştırmalı Öğrenme (CUT ve FastCUT)

CUT (Park vd., 2020) ilk önce giriş görüntü parçaları ve karşılık gelen oluşturulan görüntü parçaları arasındaki ilişkiyi öğrenerek görüntüden görüntüye dönüşüm yaparken gürültü karşılaştırmalı tahmin uygulamalar ve döngü tutarlılığına dayalı olanlardan (Zhu vd., 2017) daha üstün bir performans elde eder.

Denetimsiz görüntüden görüntüye dönüşüm için karşılaştırmalı öğrenmeyi kullanmayı daha faydalı hale getirerek, yeni karşılaştırmalı öğrenme tekniği sunulmuştur (Park vd., 2020). CUT yöntemi tek taraflı eşlemeyi iki taraflı olarak genişleterek yerleştirmeleri öğrenmede daha iyi performans göstermiştir. Önceki yöntemlerin üstesinden gelemediği mod çökmesi sorununu da karşılıklı bilgi maksimizasyonuna dayanarak ele almıştır.

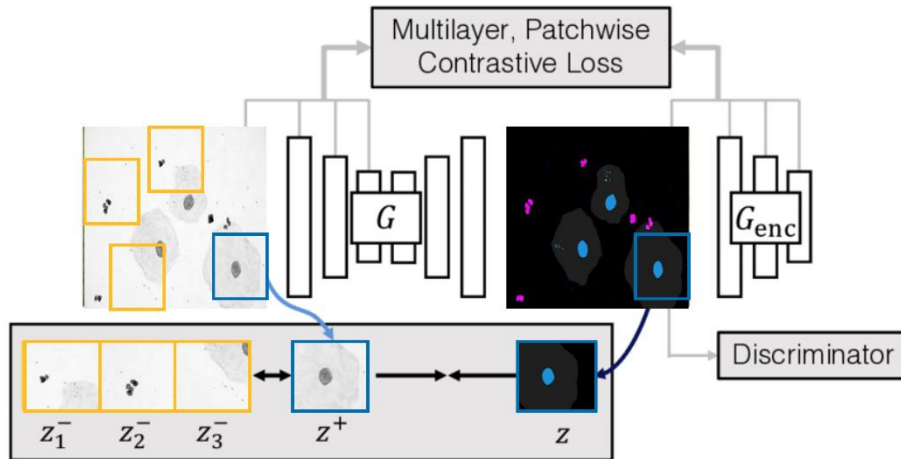
Asıl amacı oluşturulan görüntü gerçekçi olmalıdır ve giriş ve çıkış görüntülerindeki parçalar denk gelmelidir. Şekil 2’de tek taraflı dönüşüm için parça tabanlı karşılaştırmalı öğrenmede minimaks öğrenme hedefi gösterilmektedir. Oluşturulan çıkış parça görüntüsü, diğer rastgele parçalar ile karşılaştırıldığında, karşılık gelen giriş parçasına daha yakın görünmelidir. Karşılık gelen giriş ve çıkış parçaları arasındaki karşılıklı bilgiyi en üst düzeye çıkaran çok katmanlı, parça tabanlı karşılaştırmalı maliyet kullanılmaktadır. Bu, eşleştirilmemiş durumlarda tek taraflı dönüşümü mümkün kılar. Ek olarak üretici ağdan sonuç alınırken gereksiz değişim yapmasını engellemek için Y etki alanındaki görüntülerde PatchNCE yani $\mathcal{L}_{PatchNCE}(G, H, Y)$ maliyeti kullanılabilir. Bu maliyet önceki eşlenmemiş dönüşüm yöntemlerinde yaygınca kullanılan öz maliyet(identity) öğrenebilen, alana özgü bir sürümdür (Taigman vd., 2017; Zhu J.Y. vd., 2017).

FastCUT yöntemi CUT yönteminin geliştirilmiş halidir. Amaç fonksiyonu Eşitlik 3’de görüldüğü gibidir.

$$\begin{aligned}
 G^* &= \underset{G, F}{\operatorname{argmin}} \underset{D}{\operatorname{max}} \mathcal{L}(G, H, D) \\
 &= \mathcal{L}_{GAN}(G, D, X, Y) \\
 &+ \lambda_X \mathcal{L}_{PatchNCE}(G, H, X) \\
 &+ \lambda_Y \mathcal{L}_{PatchNCE}(G, H, Y)
 \end{aligned} \tag{3}$$

Burada $\lambda_X = 1$ ve $\lambda_Y = 1$ seçilerek birlikte eğitilmiştir. Regülerizasyonun yokluğunu telafi etmek için $\lambda_Y = 0$ ve $\lambda_X = 10$ seçilir. CUT yöntemi mevcut yöntemlere göre üstün performans sağladığı gösterilmiştir. Buna karşın CUT yönteminin ikinci konfigürasyonu olan ve FastCUT olarak adlandırılan yöntem CycleGAN’ın daha hızlı ve daha basit bir versiyonu olarak düşünülebileceğini bulunmuştur. Model, genellikle 5-10 maliyet ve hiper parametre kullanan son yöntemlerle karşılaştırıldığında nispeten basit bir yapıdadır. Hızlı varyant FastCUT aynı zamanda en düşük hesaplamalı eğitim maliyetiyle rekabetçi sonuçlar üretebilir (Park vd., 2020).

Şekil 2, CUT yönteminin mimarisini göstermektedir. Şekil 2’de görüntüdeki parçaların eşlenmesine maliyetine de yer verilmiştir.



Şekil 2. Parça tabanlı karşılaştırmalı öğrenme gösterimi

2.3.DCLGAN ve SimDCL

DCLGAN, iki ayrı gömülü sistem kullanarak girdi ve çıktı görüntü parçaları arasındaki özellik çıkarımını artırarak karşılıklı bilgiyi en üst seviyeye çıkarmayı amaçlar (Han vd., 2021). İkili öğrenme (Yi vd., 2017) eğitimi stabilize etmeye yardımcı olur. Karşılıklı öğrenmedeki bazı tasarım seçeneklerini yeniden gözden geçirilmiştir. Buradaki PatchNCE kaybında küçük parçaları temsil eden RGB pikselleri kaldırılmış ve daha iyi sonuçlar elde edilmiştir. Döngü tutarlılığının (Zhu vd., 2020) gereksiz olduğu gösterilmiştir. SimDCL ise DCLGAN'ın varyantıdır ve mod çökmesini önemli ölçüde önler.

DCLGAN maliyet olarak tartışmalı (Adversarial) maliyet, özdeşlik (Identity) maliyet ve CUT yönteminde de kullanılan parçalı gürültü karşılaştırmalı tahminleme (PatchNCE) maliyetini kullanır. SimDCL bu maliyetlere ek olarak benzerlik maliyetini (\mathcal{L}_{sim}) de mod çökmesini engellemek için kullanır.

DCLGAN'da oluşturulan görüntü gerçekçi olmalıdır. Giriş ve çıkış görüntüleri denk olmalıdır. Eşitlik 4'de DCLGAN'nın temel formülüne yer verilmiştir.

$$\begin{aligned}
& G^*, F^* \\
& = \underset{G, F, D_X, D_Y}{\operatorname{argminmax}} \mathcal{L}(G, F, D_X, D_Y, H_X, H_Y) \\
& = \lambda_{GAN} (\mathcal{L}_{GAN}^G(G, D_Y, X, Y) \\
& + \mathcal{L}_{GAN}^F(F, D_X, Y, X)) \quad (4) \\
& + \lambda_{NCE} \mathcal{L}_{PatchNCE_X}(G, H_X, H_Y, X) \\
& + \lambda_{NCE} \mathcal{L}_{PatchNCE_X}(F, H_X, H_Y, X) \\
& + \lambda_{idt} \mathcal{L}_{identity}(G, F)
\end{aligned}$$

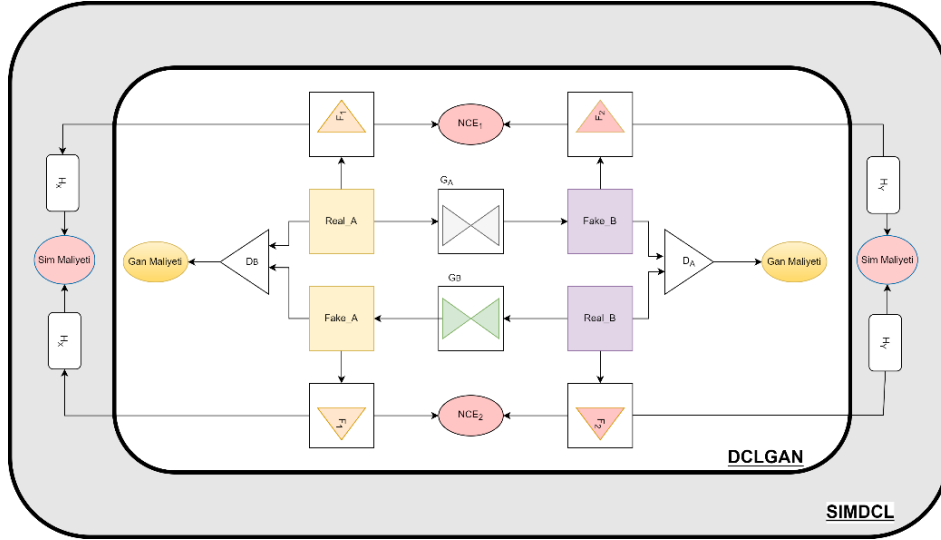
Burada $\lambda_{GAN} = 1$, $\lambda_{NCE} = 2$ ve $\lambda_{idt} = 1$ olarak alınmıştır.

DCLGAN'ın amaç fonksiyonuna benzerlik kaybı eklenmiştir ve SimDCL olarak adlandırılmıştır. Burada sim; benzerlik kısaltması ve DCL; ikili karşılaştırmalı öğrenme anlamına gelmektedir. SimDCL amaç fonksiyonu Eşitlik 5'de gösterilmiştir.

$$\begin{aligned}
& G^*, F^* \\
& = \underset{G, F, D_X, D_Y}{\operatorname{argminmax}} \mathcal{L}(G, F, D_X, D_Y, H_X, H_Y) \\
& = \lambda_{GAN} \left(\mathcal{L}_{GAN}^G(G, D_Y, X, Y) \right. \\
& + \left. \mathcal{L}_{GAN}^F(F, D_X, Y, X) \right) \quad (5) \\
& + \lambda_{NCE} \mathcal{L}_{PatchNCE_X}(G, H_X, H_Y, X) \\
& + \lambda_{NCE} \mathcal{L}_{PatchNCE_X}(F, H_X, H_Y, X) \\
& + \lambda_{sim} \mathcal{L}_{sim}(G, F, H_X, H_Y, H_1, H_2, H_3, H_4) \\
& + \lambda_{idt} \mathcal{L}_{identity}(G, F)
\end{aligned}$$

Burada $\lambda_{GAN} = 1$, $\lambda_{NCE} = 2$ ve $\lambda_{sim} = 10$ ve $\lambda_{idt} = 1$ olarak alınmıştır. Bu yapı DCLGAN'dan daha yavaş çalışır. DCLGAN'a neredeyse eşdeğer sonuçlar elde edilmiştir.

Şekil 3'te yukarıda bahsi geçen DCLGAN ve SIMDCL mimarisine yer verilmiştir.



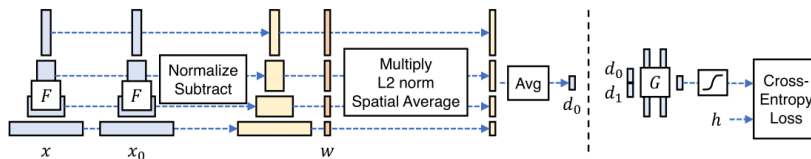
Şekil 3. DCLGAN ve SIMDCL mimarisini

2.4. Ölçüm Metrikleri

Fréchet başlangıç mesafesi (FID) (Heusel vd., 2017) ve **Kernel başlangıç mesafesi (KID)** (Bińkowski vd., 2018), ÇÜA'lerden üretilen görüntülerin kalitesini değerlendirmek için kullanılan performans ölçütüdür. FID (Heusel vd., 2017), gerçek görüntülerin dağılımını oluşturulan görüntülerin dağılımıyla karşılaştırma işlemi yapar. KID (Bińkowski vd., 2018), FID metriğine benzemektedir. Bu metrik, üretilen ve gerçek görüntü arasında karesel Maksimum Ortalama Tutarsızlık (MMD) kullanmaktadır. FID metriğine göre KID metriği ReLU aktivasyon fonksiyonu barındırdığı için avantajlıdır.

Pik Sinyal Gürültü Oranı (PSNR) (wikipedia, 2021), en yaygın kullanılan objektif görüntü kalitesi metriğidir. Aynı zamanda insanların görsel algısının karmaşık ve doğrusal olmayan etkisi nedeniyle PSNR değerleri farklılık göstermektedir.

Öğrenilmiş Algısal Görüntü Parça Benzerliği (LPIPS) (Zhang vd., 2018), iki görüntü arasındaki uzaklığı hesaplamak için kullanılan en güncel metriklerden biridir. Bu uzaklığı hesaplamak için evrişimsel sinir ağlarından yararlanır. Şekil 4'de LPIPS fonksiyonunun nasıl hesaplandığı gösterilmektedir.

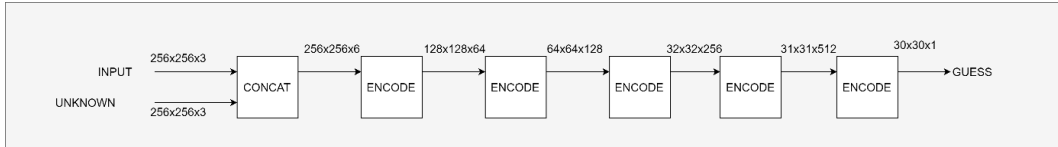


Şekil 4. LPIPS fonksiyonu. (Sol) İki görüntü parçası arasındaki uzaklık, (Sağ) İki uzaklık değerlerinden h algısal kararını tahmin etmek için yapılan küçük bir ağ eğitimi (Zhang vd., 2018)

3. Deneysel Sonuçlar

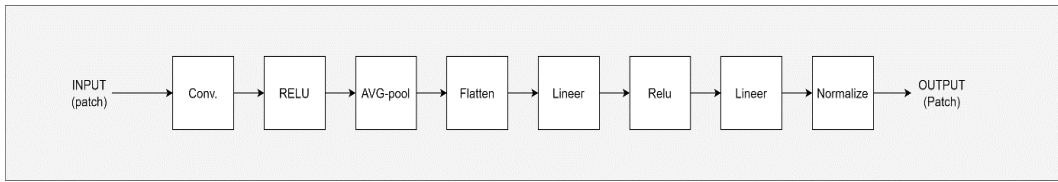
3.1. Eğitim detayları

DCLGAN ve SimDCL mimarisinde iki üretici ağ, iki özellik çıkaran katman ve iki ayrıştırıcı ağ bulunmaktadır. CycleGAN mimarisinde iki üretici ağ ve bir ayrıştırıcı ağ bulunmaktadır. CUT ve FastCUT yöntemleri bir üretici ağ, bir özellik çıkartıp gömme işlemini gerçekleştiren ağ ve bir ayrıştırıcı ağ bulundurmaktadır. Yapılan çalışmada yöntemlerin iç yapılarında bulundurduğu ağ içerikleri aynıdır. Fark yöntemlerde bulunan ağların özellikleri ve sayılarıdır. Şekil 5, ayrıştırıcı ağın iç yapısını göstermektedir. Ayrıştırıcı ağ PatchGAN mimarisini kullanmaktadır.



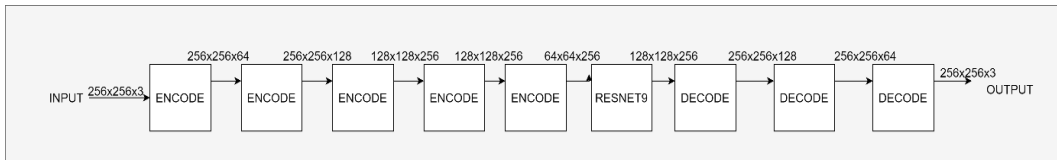
Şekil 5. Ayrıştırıcı ağ mimarisi

Şekil 6, CUT ve FastCUT yönteminde kullanılan görüntünün özelliğini çıkaran ve gömme işlemini gerçekleştiren ağ yapısına yer verilmiştir.

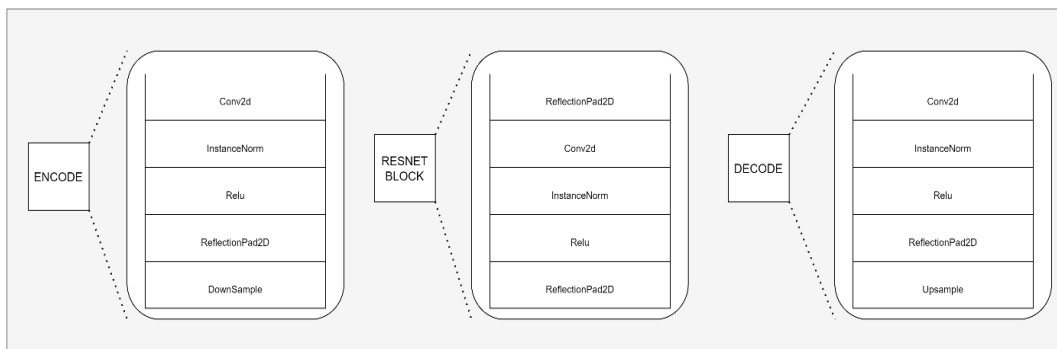


Şekil 6. Özellik çıkaran gömülü ağ mimarisi

Şekil 7'de üretici ağın genel yapısına ve ağ katmanlarına yer verilmiştir.



(a)

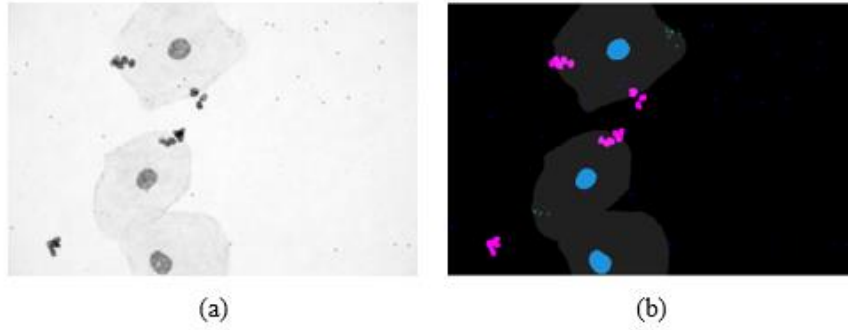


(b)

Şekil 7. Üretici ağ mimarisi (a) Üretici ağın iç yapısı (b) Üretici ağın iç yapısındaki ağ katmanlarının gösterimi

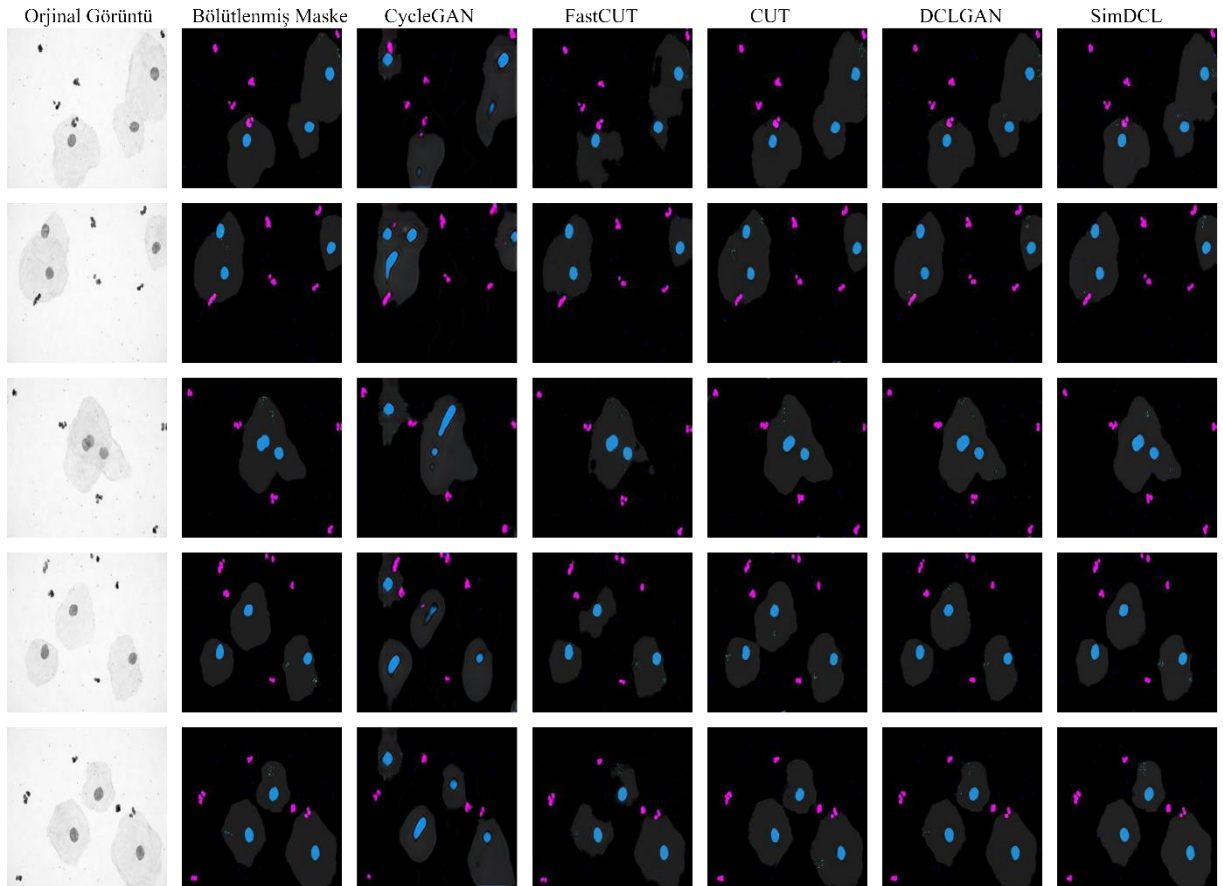
3.2. Veri Kümesi

PAPSMEAR veri kümesi, 450 tane PAPSMEAR-Maske görüntü çiftini barındırmaktadır. Veri kümesinde bulunan görseller sitoplazma, çekirdek, beyaz kan hücresi, basil ve benek nesnelere ve bu nesnelere maskelerinden oluşmaktadır. Görüntü boyutları 256x256x3 boyutundadır. Veri kümeleri örneklerine Şekil 8’de yer verilmiştir.



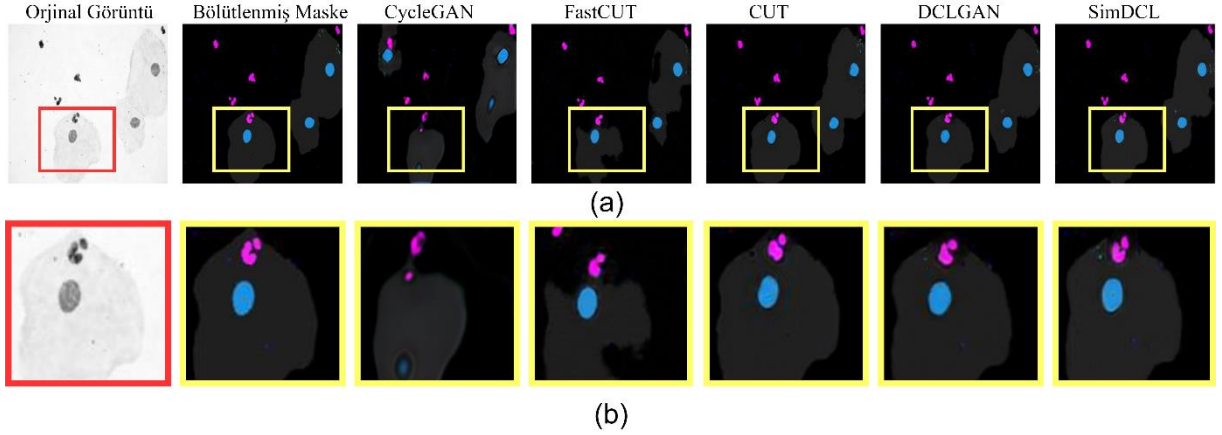
Şekil 8. PAPSMEAR veri kümesi (a) PAPSMEAR, (b) Maske

3.3. Yöntemlerin Uygulanması



Şekil 9. Yöntemlerin görsel olarak görüntü bölütleme sonuçları

Şekil 9’da PapSmear görüntü bölütlemesi yapılırken sırasıyla CycleGAN, CUT, FastCUT, DCLGAN ve SimDCL yöntemlerinin görsel sonuçlarına yer verilmiştir. Görsel çıktı sonuçlarına bakıldığında PapSmear görüntüsündeki nesne bölütlemesinin diğer yöntemler kıyaslandığında en iyi sonuçları DCLGAN ve SimDCL’in verdiği gözlemlenmektedir.



Şekil 10. Yöntemlere göre PapSmear görüntü bölütleme analizi. (a) Gerçek görüntü boyutu (256x256x3). (b) a’da gösterilen renkli kısımların yakınlaştırılmış hali

Şekil 10’da PapSmear görüntü bölütlemesi yapılırken sırasıyla CycleGAN, CUT, FastCUT, DCLGAN ve SimDCL yöntemlerinin (a) orijinal çıktılarına ve (b) yakınlaştırılmış çıktılarına yer verilmiştir. Diğer yöntemler ile kıyaslandığında bölütlenmiş maskeye sırasıyla SimDCL ve DCLGAN yönteminin daha yakın sonuç ürettiği gözlemlenmiştir. CycleGAN ve FastCUT yönteminin PapSmear görüntü bölütleme yapılırken başarım sonucunun düşük olduğu gözlemlenmiştir.

Tüm yöntemler 100 iterasyon olarak Python ortamında ve 12 GB GPU’luk bilgisayarda çalıştırılmıştır. Ölçüm metriği olarak klasik GAN yöntemlerinde kullanılan FID (Heusel vd., 2017) ve KID (Bińkowski vd., 2018) yöntemleri, CNN tabanlı LPIPS yöntemi (Zhang vd., 2018) ve geleneksel yöntem olan PSNR (wikipedia, 2021) kullanılmıştır.

Tablo 1’de ölçüm metrikleri (PSNR, LPIPS, FID, KID) kullanılarak yöntemlerin bölütlenmiş maskeye benzerlik dereceleri sunulmuştur.

Tablo 1. Görüntü benzerlik sonuçları

	PSNR	LPIPS	FID	KID
CUT	38.003	0.136	88.184	0.036
CYCLEGAN	34.101	0.333	175.451	0.137
DCLGAN	39.437	0.086	70.577	0.023
FASTCUT	36.977	0.201	128.182	0.105
SIMDCL	<u>38.920</u>	<u>0.101</u>	<u>74.203</u>	<u>0.023</u>

Tablo 1 incelendiğinde PSNR, LPIPS, FID ve KID ölçüm metriklerine göre bölütlenmiş maskeye en çok benzeyen yöntem DCLGAN’dır. İkinci sırada SimDCL yöntemi gelmektedir. CycleGAN ve FastCUT yöntemi ise ölçüm metriği sonuçlarına bakıldığında kıyaslanan yöntemler arasında başarısız olduğu gözlemlenmektedir.

4. Sonuç

Bu makalenin ana konusu görüntü bölütleme probleminde yaygın olarak kullanılan çekişmeli üretken ağlardan CycleGAN yöntemini başlangıç olarak alan yöntemlerin verimliliğinin kıyaslanmasıdır. Parça tabanlı karşılaştırmalı öğrenme yöntemlerinden olan CUT ve FastCUT yöntemi ile CycleGAN yöntemi birleştirilerek yeni ÇÜA mimarisi olarak DCLGAN ve SimDCL geliştirilmiştir. PapSmear veri kümesi bölütlemeye CycleGAN, CUT, FastCUT, DCLGAN ve SimDCL yöntemlerinin görsel bölütleme verimliliğine yer verilmiştir.

Kıyaslanan yöntemlerden görsel sonuçlar incelendiğinde en iyi sonuca sırasıyla SimDCL ve DCLGAN yöntemi olmuştur. CycleGAN ise bölütleme yaparken diğer yöntemlerin gerisinde kaldığı gözlemlenmiştir. Tablo 1'deki PSNR, FID, KID, LPIPS metrikleri ile sonuçlara bakıldığında en iyi sonuçların sırası ile DCLGAN ve SimDCL yöntemlerinin verdiği gözlemlenmiştir. Tablo 1'e göre CycleGAN ve FastCUT yönteminin kıyaslanan diğer yöntemlerden daha düşük olduğu gözlemlenmiştir. İki yöntemi (CycleGAN ve CUT) birleştiren mimari DCLGAN ve SimDCL'in PapSmear görüntü bölütlemesinde daha verimli sonuçlar elde ettiği söylenebilir.

Teşekkür

Bu çalışma İnönü Üniversitesi Bilimsel Araştırma ve Koordinasyon birimi tarafından "FDK-2021-2675" proje numarası ile finanse edilmiştir. İnönü Üniversitesi'ne teşekkürlerimizi sunarız.

Kaynaklar

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.
- Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).
- Wang, T. C., Liu, M. Y., Zhu, J. Y., Tao, A., Kautz, J., & Catanzaro, B. (2018). High-resolution image synthesis and semantic manipulation with conditional gans. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8798-8807).
- Choi, Y., Choi, M., Kim, M., Ha, J. W., Kim, S., & Choo, J. (2018). Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8789-8797).
- Azadi, S., Tschannen, M., Tzeng, E., Gelly, S., Darrell, T., & Lucic, M. (2019). Semantic bottleneck scene generation. *arXiv preprint arXiv:1911.11357*.
- Zhu, J. Y., Zhang, R., Pathak, D., Darrell, T., Efros, A. A., Wang, O., & Shechtman, E. (2017). Multimodal Image-to-Image Translation by Enforcing Bi-Cycle Consistency. In *Advances in neural information processing systems* (pp. 465-476).
- Xiong, F., Wang, Q., & Gao, Q. (2019). Consistent embedded GAN for image-to-image translation. *IEEE Access*, 7, 126651-126661.
- Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).
- Kim, T., Cha, M., Kim, H., Lee, J. K., & Kim, J. (2017, July). Learning to discover cross-domain relations with generative adversarial networks. In *International Conference on Machine Learning* (pp. 1857-1865). PMLR.
- Mejjati, Y. A., Richardt, C., Tompkin, J., Cosker, D., & Kim, K. I. (2018). Unsupervised attention-guided image to image translation. *arXiv preprint arXiv:1806.02311*.
- Liu, M. Y., Breuel, T., & Kautz, J. (2017). Unsupervised image-to-image translation networks. In *Advances in neural information processing systems* (pp. 700-708).

- Zhou, Y. F., Jiang, R. H., Wu, X., He, J. Y., Weng, S., & Peng, Q. (2019). Branchgan: Unsupervised mutual image-to-image transfer with a single encoder and dual decoders. *IEEE Transactions on Multimedia*, 21(12), 3136-3149.
- Huang, X., Liu, M. Y., Belongie, S., & Kautz, J. (2018). Multimodal unsupervised image-to-image translation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 172-189).
- Lin, J., Chen, Z., Xia, Y., Liu, S., Qin, T., & Luo, J. (2019). Exploring explicit domain supervision for latent space disentanglement in unpaired image-to-image translation. *IEEE transactions on pattern analysis and machine intelligence*, 43(4), 1254-1266.
- Park, T., Efros, A. A., Zhang, R., & Zhu, J. Y. (2020, August). Contrastive learning for unpaired image-to-image translation. In *European Conference on Computer Vision* (pp. 319-345). Springer, Cham.
- Han, J., Shoeiby, M., Petersson, L., & Armin, M. A. (2021). Dual Contrastive Learning for Unsupervised Image-to-Image Translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 746-755).
- Yurt, M., Dar, S. U., Erdem, A., Erdem, E., Oguz, K. K., & Çukur, T. (2021). Mustgan: Multi-stream generative adversarial networks for MR image synthesis. *Medical Image Analysis*, 70, 101944.
- Yao, S., Tan, J., Chen, Y., & Gu, Y. (2021). A weighted feature transfer gan for medical image synthesis. *Machine Vision and Applications*, 32(1), 1-11.
- Chabra, S. (2016). Cervical cancer preventable, treatable, but continues to kill women. *Cervical Cancer*, 1(112), 2.
- Mustafa, W. A., Halim, A., Jamlos, M. A., & Idrus, S. Z. S. (2020, April). A Review: Pap Smear Analysis Based on Image Processing Approach. In *Journal of Physics: Conference Series* (Vol. 1529, No. 2, p. 022080). IOP Publishing.
- Fekri-Ershad, S. (2019). Pap smear classification using combination of global significant value, texture statistical features and time series features. *Multimedia Tools and Applications*, 78(22), 31121-31136.
- Gautam, S., Jith, N., Sao, A. K., Bhavsar, A., & Natarajan, A. (2018). Considerations for a PAP smear image analysis system with CNN features. *arXiv preprint arXiv:1806.09025*.
- Taigman, Y., Polyak, A., & Wolf, L. (2016). Unsupervised cross-domain image generation. *arXiv preprint arXiv:1611.02200*.
- Yi, Z., Zhang, H., Tan, P., & Gong, M. (2017). Dualgan: Unsupervised dual learning for image-to-image translation. In *Proceedings of the IEEE international conference on computer vision* (pp. 2849-2857).
- Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 586-595).
- “Peak signal-to-noise ratio” https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio. Retrieved 2021.08.10
- Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. (2017). Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30.
- Bińkowski, M., Sutherland, D. J., Arbel, M., & Gretton, A. (2018). Demystifying mmd gans. *arXiv preprint arXiv:1801.01401*.