

Etkili Alt Piksel Evrişimli Sinir Ağı Süper Çözünürlük Yaklaşımı

Fatih AĞALDAY^{1*}, Ahmet ÇINAR²

¹Bilgisayar Programcılığı / Mardin Meslek Yüksekokulu, Artuklu Üniversitesi, Türkiye

²Bilgisayar Mühendisliği / Mühendislik Fakültesi, Fırat Üniversitesi, Türkiye

*(fatihagalday@gmail.com)

Özet – Tek görüntü süper çözünürlüğü, yüksek çözünürlüklü bir görüntüyü karşılık gelen düşük çözünürlüklü görüntüden kurtarmayı amaçlayan görüntü restorasyonundaki önemli içeriklerden biridir. Bu içeriklere örnek olarak kameralı izleme sisteminde insan yüzünün çözünürlüğünün düşük olması nedeniyle bazen bir kişiyi tanımak zordur. Yüz tanıma dışında, süper çözünürlüklü uygulamalar genellikle tıbbi görüntüleme ve uydu görüntüleme gibi alanlarda bulunabilir. Derin sinir ağlarına dayalı modeller tek görüntü süper çözünürlüğü için hem yeniden yapılandırma doğruluğu hem de hesaplama performansı açısından büyük başarı elde etmektedir. Bu yöntemlerde, düşük çözünürlüklü giriş görüntüsü, yeniden yapılandırmadan önce tek bir filtre, genellikle bikübik enterpolasyon kullanılarak yüksek çözünürlüklü alana yükseltilir. Süper çözünürlük işleminin yüksek çözünürlük alanında gerçekleştirildiği anlamına gelir. Bunun optimalin altında olduğunu ve hesaplama karmaşıklığı eklediğini gösteriyoruz. Bu yazıda görüntülerin gerçek zamanlı süper çözünürlük yapabilen evrişimli sinir ağını sunulmaktadır. Bu işlemleri gerçekleştirebilmek için, öznitelik haritalarının düşük çözünürlüklü görüntü uzayında çıkarıldığı yeni bir evrişimli sinir ağı mimarisi önerilmektedir. Ek olarak, son düşük çözünürlük özellik haritalarını yüksek çözünürlük çıktısına yükseltmek için bir dizi yükseltme filtresini öğrenen verimli bir alt piksel evrişim katmanı sunulmaktadır.

Anahtar Kelimeler – Süper Çözünürlük, Derin Öğrenme, Evrişimli Sinir Ağı, Alt Piksel, GANs

I. GİRİŞ

Son yıllarda, yeniden oluşturulmuş süper çözünürlüklü görüntülerin hem hesaplama performansında hem de doğruluğunda büyük başarı elde eden derin sinir ağlarına dayalı birkaç süper çözünürlük modeli vardır. Süper Çözünürlüklü Evrişimli Sinir Ağı (SRCNN), Hızlı Süper Çözünürlüklü Evrişimli Sinir Ağı (FSRCNN), Çok Derin Süper Çözünürlük (VDSR), Süper Çözünürlüklü GAN (SRGAN) üretken çekişmeli ağ ve etkili alt piksel evrişimli sinir ağı yaklaşımı ESPCN modelleri süper çözünürlük modellerine örnek olarak verilebilir. Bu yazının amacı etkili alt piksel evrişimli sinir ağı olan ESPCN modelinin süper çözünürlük yaklaşımını göstermektir.

Burada yüksek çözünürlük verileri, veri kümesindeki her orijinal görüntüyü, düşük çözünürlük verileri ise alt örneklenmiş her bir düşük çözünürlük görüntüsünü temsil etmektedir.

Yükseltme faktörünü ise r ile temsil edilmektedir. Görüntü işleme dünyasında yüksek çözünürlüklü bir görüntünün veya videonun düşük çözünürlüklü kısmından yeniden iyileştirilmesi oldukça ilgi çekmektedir. Süper çözünürlük olarak adlandırılan bu işlemler tıbbi görüntüleme [1, 2] uydu görüntüleme [3], yüz tanıma [4] ve gözetim [5] gibi birçok alanda doğrudan uygulama bulmaktadır. Genel olarak süper çözünürlük sorunu, düşük çözünürlük verilerinin yüksek çözünürlük verilerinin düşük geçişli filtrelenmiş, alt örneklenmiş ve gürültülü bir versiyonu olduğunu varsayar. Tersine çevrilemeyen düşük geçişli filtreleme ve alt örnekleme işlemleri sırasında meydana gelen yüksek frekans bilgisinin kaybı nedeniyle, oldukça yanlış oluşturulmuş bir problemdir. Süper çözünürlük işlemi, düşük çözünürlükten yüksek çözünürlük uzayına etkili bir şekilde bire çok eşlemedir. Yani birden çok çözüme

sahip olabilir ve bunlardan doğru çözümün belirlenmesi önem arz etmektedir. Birçok süper çözünürlük tekniğinin altında yatan temel varsayım, yüksek frekanslı verilerin çoğunun gereksiz olduğu ve bu nedenle düşük frekanslı bileşenlerden doğru bir şekilde yeniden oluşturulabileceğidir. Bu nedenle süper çözünürlük bir çıkarım problemidir.

Pek çok yöntem, aynı sahnenin farklı perspektiflere sahip düşük çözünürlük örnekleri olarak, yani benzersiz önceki yakın dönüşümlerle birden çok görüntünün mevcut olduğunu varsayar. Bunlar, çok görüntülü süper çözünürlük yöntemleri [6] olarak kategorize edilebilir ve kötü konumlanmış sorunu ek bilgilerle sınırlayarak ve alt örnekleme sürecini tersine çevirmeye çalışarak açık fazlalıktan yararlanır. Bununla birlikte, bu yöntemler genellikle doğruluğu sonucun kalitesini doğrudan etkileyen hesaplama açısından karmaşık görüntü kaydı ve birleştirme aşamaları gerektirir. Alternatif bir yöntem ailesi, tek görüntü süper çözünürlüklü teknikleridir [7]. Bu teknikler, tek bir düşük çözünürlüklü örnekten eksik yüksek çözünürlüklü bilgileri kurtarmak için doğal verilerde bulunan örtülü fazlalığı öğrenmeyi amaçlar.

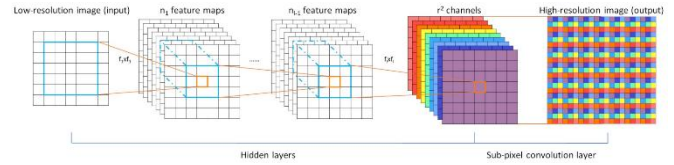
Evrişimli Sinir Ağının gelişmesiyle birlikte algoritmaların etkinliği, özellikle hesaplama ve bellek maliyetleri önem kazanmaktadır [8]. Derin ağ modellerinin doğrusal olmayan ilişkileri öğrenme esnekliğinin, daha önceki modellere kıyasla üstün yeniden yapılandırma doğruluğu elde ettiği gösterilmiştir [9, 10]. Bir düşük çözünürlük görüntüsünü yüksek çözünürlük uzayına süper çözmek için, düşük çözünürlük görüntüsünün çözünürlüğünü bir noktada yüksek çözünürlük görüntüsünün çözünürlüğüyle eşleştirmek için artırmak gerekir. Ağ içinde yükseltme, ağın son katmanı tarafından gerçekleştirilir. Bu, her düşük çözünürlük görüntüsünün doğrudan ağa beslediği ve düşük çözünürlük uzayındaki doğrusal olmayan evrişimler yoluyla özellik çıkarmanın gerçekleştiği anlamına gelir. Azaltılmış giriş çözünürlüğü nedeniyle, belirli bir bağlamsal alanı korurken aynı bilgileri entegre etmek için daha küçük bir filtre boyutunu etkili bir şekilde kullanabilmektedir.

II. MATERYAL VE YÖNTEM

Tek görüntü süper çözünürlüğün görevi, orijinal yüksek çözünürlüklü görüntüye karşılık küçültülmüş düşük çözünürlüklü bir görüntü verildiğinde, yüksek çözünürlüklü bir görüntünün

süper çözünürlüğünü tahmin etmektir. Alt örnekleme işlemi belirleyicidir ve yüksek çözünürlükten düşük çözünürlük üretmek için önce yüksek çözünürlüğü bir Gauss filtresi kullanılarak evriştirilir ve böylece kameranın nokta yayma işlevini simüle eder sonrasında görüntüyü r faktörü kadar alt örnekleme yapılır. Yükseltme oranı olarak r faktörü kullanılır. Genel olarak, hem düşük çözünürlük ve hem de yüksek çözünürlük görüntüler C renk kanallarına sahip olabilir, sırasıyla $H \times W \times C$ ve $rH \times rW \times C$ boyutunda gerçek değerli çok boyutlu veri olarak temsil edilirler.

Süper çözünürlük problemini çözmek için, [11]'de önerilen SRCNN, düşük çözünürlük yerine yükseltilmiş ve enterpolasyonlu düşük çözünürlüklü bir versiyondan kurtarır. Süper çözünürlük görüntüyü iyileştirmek için 3 katmanlı evrişimli bir ağ kullanılır. Bu bölümde, düşük çözünürlük görüntüyü ağa beslemeden önce yükseltmekten kaçınmak için Şekil 1'de gösterildiği gibi yeni bir ağ mimarisi önerilmektedir. Mimaride, önce l katmanlı evrişimli sinir ağını doğrudan düşük çözünürlük görüntüsüne uyguluyor ve ardından süper çözünürlük üretmek için düşük çözünürlük özellik haritalarını yükselten bir alt piksel evrişim katmanı uygulanmaktadır.



Şekil 1. ESPCN Modelinin süper çözünürlük adımları

A. Ters Evrişim Katmanı

Bir ters evrişim katmanının eklenmesi, maksimum havuzlama ve diğer görüntü aşağı örnekleme katmanlarından çözünürlüğü kurtarmak için popüler bir seçimdir. Bu yaklaşım, katman aktivasyonlarının görselleştirilmesinde ve ağdan yüksek seviyeli özellikler kullanılarak semantik segmentasyonların oluşturulmasında başarılı bir şekilde kullanılmıştır [12]. Önerilen ters evrişim katmanı, her bir giriş pikselinin, r adımı ile eleman bazında bir filtre ile çarpılması olarak görülebilir ve geriye doğru evrişim olarak da bilinen sonuçtaki çıktı pencerelerinin toplamları olarak görülebilir [12].

B. Verimli alt piksel evrişim katmanı

Düşük çözünürlük görüntüsünü yükseltmenin bir diğer yolu, yukarıda bahsedildiği gibi düşük

çözünürlük uzayında $\frac{1}{r}$ adım ile evrişimdir. Bu, düşük çözünürlük uzayından yüksek çözünürlük uzayına interpolasyon, delik açma [13] veya havuzdan çıkarma [14] yoluyla ve ardından bir konvolüsyon ile yüksek çözünürlük alanında 1 adım ile uygulanabilir. Evrişim yüksek çözünürlük uzayında gerçekleştiğinden, bu uygulamalar hesaplama maliyetini r^2 faktörü kadar artırır.

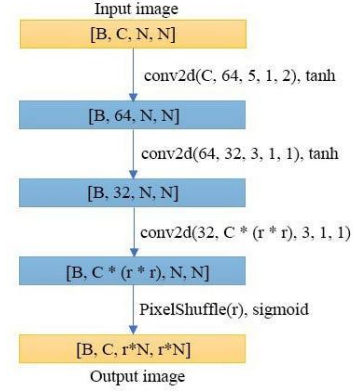
C. CNN Yaklaşımları ile ESPCN Arasındaki Fark

SRCNN, FSRCNN ve VDSR gibi evrişimli sinir ağına dayalı yaklaşımların bazı sakıncaları vardır. CNN (Evrişimli Sinir Ağı) yaklaşımlarının, çift kübik enterpolasyon gibi düşük çözünürlük görüntüsünü üst örnekleme için enterpolasyon yöntemlerini kullanması gerekir. Ağın ilk katmanında veya öncesinde çözünürlüğü artırır. Başka bir ifade ile CNN yaklaşımı, hesaplama karmaşıklığını ve bellek maliyetini artıracak olan üst örnekleme düşük çözünürlük görüntüsüne doğrudan evrişimli sinir ağını uygular. Bu sorunu çözmek için, CNN ağına verimli bir alt piksel evrişim katmanı eklemek için yeni yaklaşım Etkili Alt Piksel Evrişimli Sinir Ağı (ESPCN) önerilmektedir. ESPCN, ağın en ucundaki çözünürlüğü artırır. ESPCN modelin yükseltme adımı son katman tarafından gerçekleştirilir, bu da daha küçük boyutlu düşük çözünürlük görüntüsünün doğrudan ağa beslendiği için enterpolasyon yöntemini kullanmaya gerek yoktur. Ağı beslenmeden önce bir enterpolasyon filtresi yükseltmesine kıyasla daha iyi bir düşük çözünürlük ve yüksek çözünürlük eşleşmesi öğrenme yeteneğine sahiptir. Girdi görüntü boyutunun küçültülmesi nedeniyle, özellikleri çıkarmak için daha küçük bir filtre boyutu kullanılmaktadır. Hesaplama karmaşıklığı ve bellek maliyeti azaltılırken verimlilik büyük ölçüde artırılabilir. Bu nedenle ESPCN, gerçek zamanlı olarak HD videoların süper çözünürlüğü için ideal bir model haline gelmektedir.

D. Ağ Yapısı

Süper çözünürlük modeli genel olarak girdi verilerinin düşük çözünürlüğe sahip bulanık ve gürültülü bir görüntü olduğunu varsayar. Düşük çözünürlük görüntüleri, veri kümelerinden yüksek çözünürlük görüntüleri üzerinde bir alt örnekleme işlemi gerçekleştirilerek oluşturulabilir. Ve modelin çıktısı belirtilen üst ölçek faktörü ile yeniden yapılandırılmış süper çözünürlük görüntü olacaktır.

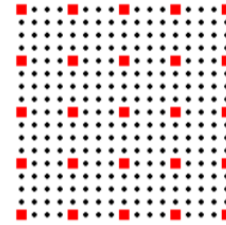
ESPCN modeli yapısı Şekil 2'de gösterilebilir. Ağ için L katmanları olduğunu varsayalım, ilk $L-1$ katmanları, giriş düşük çözünürlük görüntülerinin özellik haritalarını elde eden evrişimli katmanlardır. Ve son katman, çıktı görüntü boyutunu belirtilen bir üst ölçek faktörü ile kurtarmak için verimli alt piksel evrişimli katmandır. Şekil 2'de ESPCN modelin 3 katmana sahip ağ yapısı gösterilmektedir. Şekilde belirtilen ilk kısmında B batch size, C Channel size, N image size, R upscale faktör olarak verilmiştir.



Şekil 2. ESPCN Modeli

E. Alt piksel Evrişimi

Modelde önerilen en önemli kavramlardan biri, piksel karıştırma olarak da bilinen alt piksel evrişimidir. Altpiksel evrişimini anlamadan önce, alt piksel kavramının iyi anlaşılması gerekmektedir. Kamera görüntüleme sisteminde, kamera tarafından elde edilen görüntü verileri bir tür ayrık işleme yöntemi ile işlenmiştir. Işık sensörünün sınırlaması nedeniyle, görüntüler orijinal piksel çözünürlüğüyle sınırlıdır, başka bir deyişle görüntülerdeki her piksel, gerçek dünyada küçük bir renk alanını temsil eder. Gördüğümüz dijital görüntüde pikseller birbirine bağlıyken, mikroskobik dünyada iki fiziksel piksel arasında çok sayıda küçük piksel vardır. Bu küçük piksellere alt piksel denir

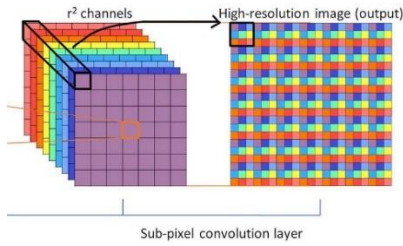


Şekil 3. Alt Piksellerin Görselleştirilmesi

Şekil 3'te dört küçük kırmızı kareyle çevrili her bir kare alan kameranın görüntüleme düzlemindeki piksel ve siyah noktalar olarak görünen alt pikseller gösterilmektedir. Alt piksellerin doğruluğu, bitişik

pikseller arasındaki enterpolasyona bağlı olarak ayarlanabilir. Bu şekilde, küçük kare alanlardan büyük kare alanlara eşleme, alt piksel enterpolasyonu yoluyla uygulanabilir.

Bu teoriye dayanarak, yüksek çözünürlüklü görüntüler elde etmek için super çözünürlük modelinde alt piksel evrişim yöntemi kullanılabilir. Ters evrişim işleminde, görüntüleri sıfırlarla doldurulur ve ardından sonuç için kötü olabilecek evrişimi yapılır. Ağın son katmanında piksel karıştırma gerçekleştirirken, düşük çözünürlük görüntüsünü kurtarmak için dolgu işlemine ihtiyaç duymaz. Şekil 4'te gösterildiği gibi çok kanallı özellik haritalarındaki her bir pikseli çıkış görüntüsünde bir $r \times r$ kare alana birleştirir. Böylece, özellik haritalarındaki her piksel, oluşturulan çıktı görüntüsündeki alt piksele eşdeğerdir.



Şekil 4. Piksel Karıştırma İşlemi

Alt piksel evrişimi iki temel süreci içerir: genel bir evrişim işlemi ve ardından piksellerin yeniden düzenlenmesi. Toplam piksel sayısının elde edilecek yüksek çözünürlük görüntüsü ile tutarlı olması için son katmanın çıkış kanalı $C \times r \times r$ olmalıdır. ESPCN ağında, enterpolasyon yöntemi dolaylı olarak evrişimli katmanlarda bulunur, ağ tarafından otomatik olarak öğrenilebilir. Konvolüsyon işlemleri daha küçük boyutlu düşük çözünürlük görüntüleri üzerinde uygulandığı için verimlilik çok daha yüksektir.

F. Kayıp Fonksiyonu

Modelde düşük çözünürlük görüntüleri, veri kümesindeki yüksek çözünürlük görüntülerinin alt örnekleme ile oluşturulur. Eğitim sürecinde, orijinal yüksek çözünürlük görüntüleri yer gerçeği verileri olacaktır. Ortalama kare hatası (MSE), oluşturulan super çözünürlük görüntüleri ile yer gerçeği yüksek çözünürlük görüntüleri arasındaki farkı ölçmek için kullanılır. Ağın piksel bazında MSE kayıp işlevi aşağıdaki şekildedir.

$$\ell(W_{1:L}, b_{1:L}) = \frac{1}{r^2 HW} \sum_{x=1}^{rH} \sum_{y=1}^{rW} (\mathbf{I}_{x,y}^{HR} - f_{x,y}^L(\mathbf{I}^{LR}))^2$$

Şekil 6. ESPCN'nin Kayıp Fonksiyonu

III. SONUÇ

ESPCN, evrişimli sinir ağı super çözünürlük modelinin yükseltilmiş bir versiyonu olarak görülebilir. ESPCN'de ağ, çeşitli evrişimli katmanların ve bir alt piksel evrişimli katmanın bir kombinasyonudur, düşük çözünürlük görüntüsü son piksel karıştırma aşamasında yükseltilir. Bu nedenle, ESPCN'nin büyük bir avantajı, daha yüksek bir hesaplama hızına sahip olmasıdır. Bu arada, diğer CNN super çözünürlük modellerine kıyasla yüksek çözünürlük görüntülerini ve videolarını kurtarma konusunda iyi bir performansa sahiptir.

KAYNAKLAR

- [1] S. Peled and Y. Yeshurun. Superresolution in MRI: application to human white matter fiber tract visualization by diffusion tensor imaging. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 45(1):29–35, 2001. 1
- [2] W. Shi, J. Caballero, C. Ledig, X. Zhuang, W. Bai, K. Bhatia, A. Marvao, T. Dawes, D. O'Regan, and D. Rueckert. Cardiac image super-resolution with global correspondence using multi-atlas patchmatch. In K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, editors, *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 8151 of LNCS, pages 9–16. 2013. 1
- [3] M. W. Thornton, P. M. Atkinson, and D. a. Holland. Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping. *International Journal of Remote Sensing*, 27(3):473–491, 2006. 1
- [4] B. K. Gunturk, A. U. Batur, Y. Altunbasak, M. H. Hayes, and R. M. Mersereau. Eigenface-domain super-resolution for face recognition. *IEEE Transactions on Image Processing*, 12(5):597–606, 2003. 1
- [5] L. Zhang, H. Zhang, H. Shen, and P. Li. A super-resolution reconstruction algorithm for surveillance images. *Signal Processing*, 90(3):848–859, 2010. 1
- [6] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344, 2004. 1
- [7] C.-Y. Yang, C. Ma, and M.-H. Yang. Single-image super-resolution: A benchmark. In *European Conference on Computer Vision (ECCV)*, pages 372–386. Springer, 2014. 1, 2
- [8] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *CVPR 2015*, 2015. 2, 7
- [9] Z. Wang, D. Liu, J. Yang, W. Han, and T. Huang. Deeply improved sparse coding for image super-resolution. *arXiv preprint arXiv:1507.08905*, 2015. 2, 3, 4

- [10] S. Schulter, C. Leistner, and H. Bischof. Fast and accurate image upscaling with super-resolution forests. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 3791–3799, 2015. 2, 3, 4, 6, 7
- [11] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2015. 2, 3, 4, 5, 6, 7, 8
- [12] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. arXiv preprint arXiv:1411.4038, 2014. 3, 4
- [13] C. Osendorfer, H. Soyer, and P. van der Smagt. Image superresolution with fast approximate convolutional sparse coding. In Neural Information Processing, pages 250–257. Springer, 2014. 2
- [14] M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks. In Computer Vision–ECCV 2014, pages 818–833. Springer, 2014. 2, 3, 4