



MRG Yüzeyi Üzerinden Kafa Pozu Tahmini için Veri Artırımı

Data Augmentation for Head Pose Estimation From MRI Surface

Ahmet Emin Yetkin¹, Andaç Hamamcı¹

¹Biyomedikal Mühendisliği Bölümü, Yeditepe Üniversitesi, İstanbul, Türkiye
{ahmetemin.yetkin@std.yeditepe.edu.tr}

Özetçe—Bu çalışmada, hastanın düzlemsel görüntüsü ile daha önceden alınmış baş MR görüntüleri arasındaki bağıl pozun kestirilmesi için konvolüsyonel sinir ağı (CNN) tabanlı bir yöntem sunulmuştur. Görüntü kümesinin boyutlanması ve ışık yönlülüğünün CNN sınıflandırıcının sağlamlığı üzerine etkisi incelenmiştir. Sonuçlar öğrenme işleminin farklı boyutlarda görüntüler içerdiği veri setlerinden olduğu zaman görüntünün boyutundaki değişimlere karşı daha sağlam olduğunu gösterdi. Diğer yandan, yüzdeki aydınlatmada açısız ışığın şiddetinin homojen ışığa daha baskın olmasıyla sınıflandırıcının sağlamlığının da arttığı görülmüştür.

Anahtar Kelimeler — Kafa Poz Tahmini; MRG yüzeyi; Derin öğrenme

Abstract—In this study, a convolutional neural network (CNN) based method to estimate the relative pose between a planar image of the patient and his MR volume is presented. The effect of data augmentation by object magnification and directional lightning on the robustness of the classifier is investigated. Results demonstrate that training from randomly scaled images were more robust to variations on image size. On the other hand, the robustness of the classifier is increased when the angular light become more dominant than homogenous light.

Keywords — Head Pose Estimation; MRI Surface; Deep Learning.

I. GİRİŞ

Vücudun konum ve yönelim bilgilerinin cerrahi/girişimsel işlem esnasında saptanması görüntü rehberliğinde tedavi (IGT) uygulamalarında sıklıkla karşılaşılan bir gereksinimdir. Bu amaçla geliştirilen klinik sistemler çoğunlukla hasta vücuduna sabitlenen işaretlerin optik, elektromanyetik veya ultrasonografik olarak takip edilmesine dayanmaktadır [1].

Artırılmış gerçeklik başta olmak üzere poz tahmini diğer birçok alanda da önemli bir role sahip, güncel bir problemdir. 2 boyutlu görüntüden yapılan kafa pozu tahmini aydınlatma koşullarına, niteliklerin eksik veya

beklenenden farklı gözükmesine hassastır[2]. Nitelik bazı çalışan algoritmalar her ne kadar yüksek doğrulukta çalışabiliyor olsa da objenin tanımlanmasında kıstas olarak belirlenen niteliklerin farklı açılardan, farklı ortam koşullarında farklı görüntülenmesi veya hiç görüntülenmemesi bu tür algoritmaların çalışma koşullarını kısıtlamaktadır.[3] Kişinin gözlük şapka gibi aksesuar kullanması, sergilediği yüz ifadeleri bile sonucu değiştirici etkenler arasında yer almaktadır.

Poz tahmini için yapay zekâ kullanımı göreceli olarak daha yeni bir yöntemdir. Mevcut pozu veri setindeki etiketlenmiş örneklerle karşılaştırarak denetimli öğrenme yöntemi olan geri yayılımla ağ yapısındaki ağırlıkları yeniden düzenleyerek veri setinde olmayan bir pozu tahmin edebilir.[4] Bu çalışmada kafa pozunun kestirimi için gözetimli derin öğrenme türü olan konvolüsyonel sinir ağları temelli bir yöntem önerilmiştir.

Standart bir yapay sinir ağı, birbirine bağlı nöron dediğimiz işleyicilerden oluşmaktadır. Her biri kendilerine atandığı değerler doğrultusunda aktif olurlar. Yapay sinir ağının mevcut koşullara göre sergilediği davranışın başarıya ulaşması nöronlara atanan değerlerin uygunluğu ve nöronların bağlantıları ile alakalıdır. Derin öğrenme, temelde çoklu katmanlı yapay sinir ağlarına dayanan algoritmalarla oluşan makine öğrenme türüdür. Nöronların ağırlıklarının geriye dönük belirlenmesi doğruluğu daha yüksek algoritmalar elde etmeye yaramaktadır.[5]

Derin öğrenme makine görüşüne ve makine öğrenmesine ait modülerin birleştiği bir alandır. Geleneksel makine öğrenme teknikleri veriyi ham şeklinde işlemeyle sınırlıdır. Bu yöntemlerin istenilen sonuçları vermesi için görüntüde öğrenilmesi istenilen niteliklerin özenle ortaya çıkarılması ve ham veri haline dönüştürülmesi gerekiyordu. Derin öğrenmeyi önceki makine öğrenme türlerinden avantajlı kılan, önem taşıyan nitelikleri insan müdahalesi gereksizdir çıkarılabilmesidir. Derin öğrenme yöntemleri görüntünün en basit halinden başlayarak farklı



Görüntü İşleme 1

1. Gün / 27 Ekim 2016, Perşembe

dönüşümler kullanarak görüntünün içinde yer alan daha karmaşık yapıları öğrenme yetisine sahiptir. Örnek olarak, Piksel dizisi halinde gelen bir görüntüde, ilk katman kenarları ve konumlarını, ikinci katman genel motifleri, sonraki katman bu motiflerin kombinasyonlarını derleyerek benzer nesnelerin varlığını araştırabilir. [6]

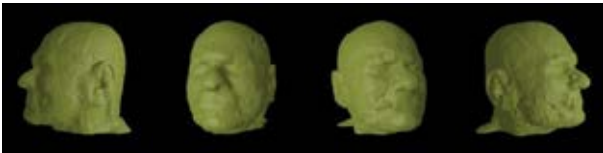
Bu çalışmanın amacı, hastanın optik kamerayla alınmış görüntüsü ile daha önceden alınmış baş MR görüntüleri arasındaki bağıl pozun kestirilmesidir. Bu amaçla, CNN, etiketlenmiş poz görüntüleriyle eğitilmiş ve sonrasında veri setinde yer almayan bir örnek için poz kestirmesi istenmiştir. Görüntünün kümesinin boyutlanması ve ışık yönlülüğünün CNN sınıflandırıcının sağlamlığı üzerine etkisi incelenmiştir.

II. YÖNTEM

Çalışmada Yeditepe Üniversitesi Hastanesinde sağlıklı bir denekten 3 Boyutlu gradyan-eko sekansı kullanılarak 0.8x0.8x1.0 mm çözünürlükte alınan MR hacmi kullanılmıştır. Görüntüleme hacmi 240 mm kare FOV ve 200 aralıksız kesit olarak tüm kafayı kapsayacak şekilde belirlenmiştir.

Kafa yüzey modeli oluşturmak için, öncelikle MR görüntülerinden kafa hacmi ITK-SNAP v3.4 programı ile bölütlenmiştir (<http://www.itksnap.org/>). Bölütleme işlemi 3 boyutlu bölge tabanlı aktif kontur yöntemi kullanılmıştır. Elde edilen etiket haritası, 3D Slicer v4.0 programı kullanılarak sayısal yüzey modeline dönüştürülmüştür (<http://www.slicer.org/>).

Oluşturulan 3B model render işlemlerini yapmak üzere 3D Studio MAX programına aktarılmıştır. Farklı koşullar altında görselleştirmeler 3D Studio MAX programının Phytion arayüzü kullanılarak oluşturulmuştur. Kameranın düşey eksen çevresinde (sağ-sol) yarım daire üzerinde 10° aralıklarla 180° döndürülmesiyle Şekil 1'de örnekleri verilen 18 farklı açıdan görünüm oluşturulmuş ve 18 farklı sınıf olarak etiketlenmiştir. Modelin aydınlatılması için, homojen bir etki gösteren gökyüzü ışığı ve açısız ışık birlikte kullanılmıştır. Açısız ışık, her bir kamera açısı için, yine düşey eksen çevresinde ve kamera açısını merkeze alarak -90° ile +90° arasında 10°lik adımlarla gezdirilmiştir. Böylelikle 18 farklı poz sınıfının her biri için 18 farklı ışık açısıyla toplam 324 görüntü, siyah arka plan üzerine ve 256x256 piksel boyutlarında render edilmiştir.



Şekil 1. Farklı ışık ve kamera açılarına sahip render sonucu elde edilen görüntüler.

Bu görüntülerden 260 tanesi eğitim ve 64 tanesi test kümesi olarak ayrılmış ve konvolüsyonel sinir ağı (CNN)

sınıflandırıcının eğitilmesi için kullanılmıştır. CNN işlemleri için, "Berkeley Vision and Learning Center" tarafından geliştirilen Caffe derin öğrenme iskeleti (<http://caffe.berkeleyvision.org/>) Windows işletim sistemi altında derlenmiştir. Ağ yapısı olarak "caffenet" kullanılmıştır. Doğrulama işlemleri için sınıflandırma MATLAB arayüzünde gerçekleştirilmiştir.

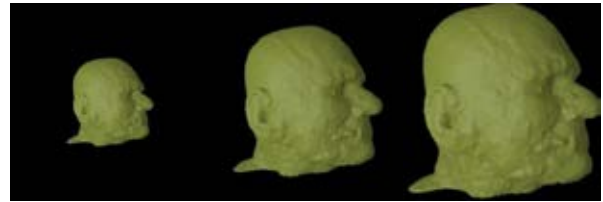
Bu çalışmada eğitim kümesinde görüntü boyutu seçimi ve uygulanan ışığın yönlülüğünün sınıflandırıcının sağlamlığı üzerindeki etkisini saptamak amacıyla iki farklı deney sunulmuştur.

A. Görüntü Boyutu

Bu bölümde, CNN sınıflandırıcının eğitildiği boyuttan farklı boyuttaki örnekler üzerindeki performansı değerlendirilmiş ve 3 farklı şekilde oluşturulmuş eğitim kümesinin sağlamlıkları karşılaştırılmıştır.

Bu amaçla, orijinal eğitim kümesinden seçilen 3 örnek, şekil 2'de gösterildiği gibi, 0.3 ile 1.3 arasında değişen oranlarda boyutlandırılmış ve her bir boyutta sınıflandırıcının doğru etiket için verdiği olasılık değerlerinin ortalaması hesaplanmıştır. İşlem, aşağıda verilen 3 farklı veri kümesiyle eğitilen sınıflandırıcılar için tekrarlanmıştır:

- 1- Orijinal (sabit) boyuttaki 260 örnekten oluşan eğitim kümesi.
- 2- Orijinal eğitim kümesinin her bir örneği 0.8 ile 1.2 oranı arasında rastgele bir oranda boyutlandırılmasıyla elde edilmiş 260 örnekten oluşan eğitim kümesi.
- 3- Madde 2 deki işlemin iki kere tekrarlanmasıyla elde edilmiş, 520 örnekten oluşan eğitim kümesi.



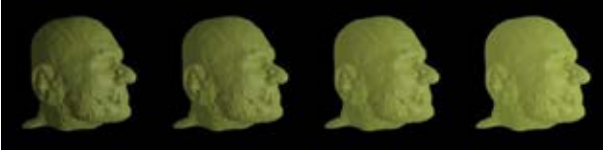
Şekil 2. Sırasıyla 0.3 (sol), 1.0 (orta) ve 1.3 (sağ) oranlarında boyutlandırılmayla oluşturulmuş örnek görüntüler

B. Aydınlatma Yönlülüğü

Aydınlatma yönlülüğünün sınıflandırıcının sağlamlığına etkisini ölçmek amacıyla, farklı ağırlıklarda homojen ışıklandırma uygulanan 4 farklı eğitim kümesi oluşturulmuş ve her bir kümenin diğer kümelerin örnekleri üzerindeki çapraz sınıflandırma performansları belirlenmiştir. Homojen aydınlatma sağlayan gökyüzü ışığının bağıl şiddetinin 0.05, 0.1, 0.2 ve 0.3 olarak atanmasıyla elde edilen 4 veri kümesinden örnek görüntüler Şekil 3'de verilmiştir.

Görüntü İşleme 1

1. Gün / 27 Ekim 2016, Perşembe



Şekil 3. Aynı kamera ve ışık kaynağı açısına sahip, gökyüzü ışığı bağlı şiddeti soldan sağa sırasıyla 0.05, 0.1, 0.2 ve 0.3 olan görüntüler.

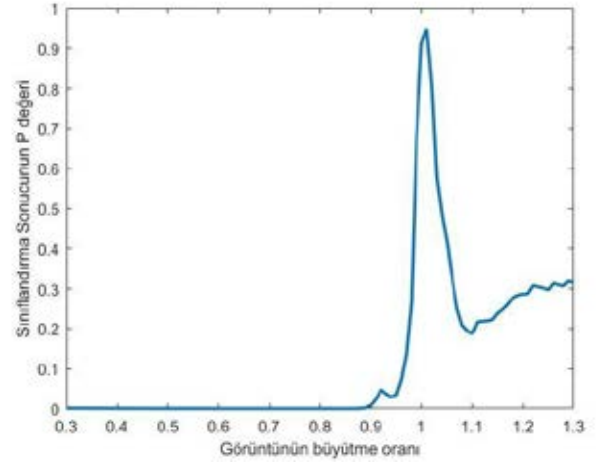
III. SONUÇ

Farklı aydınlatma koşullarında eğitilmiş modellerin farklı aydınlatma koşullarındaki sonuçları tablo.1'deki gibidir.

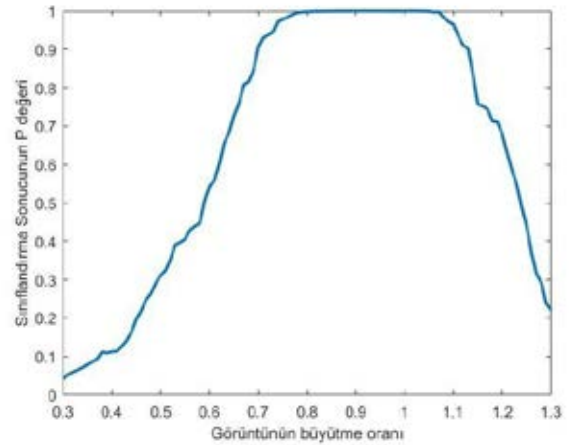
<i>Model</i> <i>Test</i>	0.05	0.1	0.2	0.3
0.05	0.97	0.54	0.48	0.26
0.1	0.99	0.81	0.66	0.34
0.2	0.88	0.95	0.94	0.59
0.3	0.69	0.94	0.98	0.8

Tablo 1. Modellerin farklı aydınlatma koşulları altındaki p değeri sonuçları. 1. Satır ve 1. sütunda yer alan sayılar gökyüzü ışığı katsayısıdır.

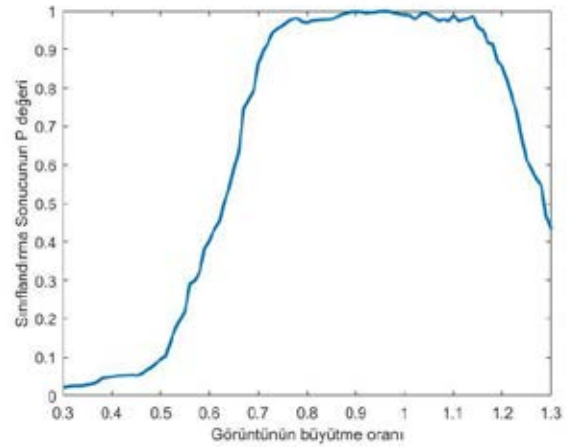
Görüntünün boyutundaki değişime bağlı olarak aynı boyuttaki 260 görüntüden eğitilmiş modelin sonucu şekil4'te, rastgele boyuttaki 260 görüntüden eğitilmiş modelin sonucu şekil 5te, rastgele boyuttaki 520 görüntüden eğitilmiş modelin sonucu şekil6'da yer almaktadır.



Şekil 4. Aynı büyüklükteki 260 görüntüyle eğitilen modelin farklı büyüklükteki görüntülere karşı sonucu.



Şekil 5. Rastgele büyüklükteki 260 görüntüyle eğitilen modelin farklı büyüklükteki görüntülere karşı sonucu



Şekil 6. Rastgele büyüklükteki 520 görüntüyle eğitilen modelin farklı büyüklükteki görüntülere karşı sonucu



Görüntü İşleme 1

1. Gün / 27 Ekim 2016, Perşembe

IV. TARTIŞMA

Bu çalışmada, MRG verisini kullanarak elde ettiğimiz yüz ve kafa modelinin yalpa hareketi pozu üzerinde görüntü boyutu. Farklı koşulların etkilerini ölçmek için veri artırımını inceledik. Bu çalışmada, kafa poz tahmininin sadece yalpa hareketi üzerinden, yani 1 serbestlik açısı üzerinden incelenmesi yapılmıştır. Kafa pozisyonu ve yönü sabit kalıp kameranın hareketi söz konusudur. Kameranın yan profilden görüntü alması ile kafanın yana dönmesi aynı şeyler değildir.

Veri kümemiz MR yüzeyinden modellenen görüntülerden oluşmaktadır. Bu veri setinden eğitilen modeli gerçek insan görüntüsünde kafa poz modelinde kullanabilmek için bir takım işlemlerden geçirilmesi gerekmektedir. Render işleminde amaç, cismin aynı açıdan alınmış görüntüsünü değiştirebilecek ve/ veya farklı açıdan alınmış iki farklı görüntüyü birbirine benzetebilecek gerçek ortam şartlarındaki değişkenlerin taklit edilerek yapay zekâyı bu koşullara karşı eğitmektir.

Aydınlatmanın temel rolü, yüzde yer alan niteliklerin bazılarını aydınlatarak bazılarını gölgede bırakarak ön plana çıkarmasıdır. Söz konusu etkiyi ölçmek için ortam açısal ışıkla beraber gökyüzü ışığı da ekledik. Açısal ışığın şiddetini ve açı aralığını sabit tutarak Gökyüzü ışığının çarpanını 0.05, 0.1, 0.2, 0.3 yaparak farklı setler elde ettik. Her bir seti ayrı öğrenme işlemine tabi tuttuk. Sonrasında modelleri farklı ışık koşulları altında nasıl sonuç verdiklerini inceledik. Tablol'deki en düşük performanslar göz önünde bulundurulduğunda homojen aydınlatma sağlayan gökyüzü ışığının fazla olduğu veri setleriyle eğitilmiş modeller diğer ışık şartlarına sahip görüntülerin pozunu tespit etmekte düşük başarı göstermektedir.

Görüntü boyutunun önemi ise, kameranın her zaman kafa pozunu tek bir büyüklükte yakalayamamasından kaynaklanmaktadır. Eğer poz tahmini algoritması aynı görüntünün makul oranlarda büyütülmüş veya küçültülmüş hallerini farklı pozlara benzetiyor veya hiç bir poza benzetemiyorsa başarısız olmuş demektir. Rastgele büyüklüklerdeki görüntülerle eğitilen modelin bu koşula direnç kazandığı şekil 4 ve şekil 5 mukayesesinde açıkça görülmektedir. Buna ek olarak şekil 5'e baktığımızda küçültülmüş görüntülerde büyütülen görüntülere oranla daha iyi sonuç aldığı görülmektedir. Veri setini iki katına çıkardığımızda sonucun daha iyileştiğini gördük.

KAYNAKÇA

- [1] Tokuda J, et al. "OpenIGTLink: an open network protocol for image-guided therapy environment", MRCAS. 5(4):423-434, 2009.
- [2]Breitenstein, Michael D., et al. "Real-time face pose estimation from single range images." Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008.
- [3]Hu, Zhiang. "Real-Time Head Pose Estimation with Convolutional Neural Networks."Lee, K.-F., *Automatic Speech Recognition: The Development of the SPHINX SYSTEM*, Kluwer Academic Publishers, Boston, 1989.
- [4]Murphy-Chutorian, Erik, and Mohan Manubhai Trivedi. "Head pose estimation in computer vision: A survey." IEEE transactions on pattern analysis and machine intelligence 31.4 (2009): 607-626.
- [5]Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." Neural Networks 61 (2015): 85-117.
- [6] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." Nature 521.7553 (2015): 436-444.