

# Derin Öğrenme Kullanarak Gerçek Dünya Doku Görüntülerinin Zorlu Koşullarda Tanınması

## Recognition of Real-World Texture Images Under Challenging Conditions With Deep Learning

Özal Yıldırım<sup>1</sup>, Ayşegül Uçar<sup>2</sup>, Ulaş Baran Baloğlu<sup>1</sup>

<sup>1</sup>Bilgisayar Mühendisliği Bölümü, Munzur Üniversitesi, Tunceli, Türkiye  
{oyildirim, ubbaloglu}@munzur.edu.tr

<sup>2</sup>Mekatronik Mühendisliği Bölümü, Fırat Üniversitesi, Elazığ, Türkiye  
agulucur@firat.edu.tr

**Özetçe**—Gerçek dünyadan elde edilen görüntüler üzerinde çeşitli kalite bozulmaları meydana gelebilmektedir. Örneğin hareket halindeki bir nesnenin görüntülenmesi veya bir ortamın hareket halindeyken görüntülenmesi görüntü üzerinde hareket izi etkisi oluşturmaktadır. Bu tür şartlarda elde edilen görüntü verileri üzerinde tanıma işlemi gerçekleştiren uzman sistemlerin yüksek tanıma performansı sağlaması önemli bir çalışma konusudur. Bu makale çalışmada, zorlu şartlar altında görüntülerin yüksek başarımla sınıflandırılması için Konvolüsyonel Sinir Ağı (KSA) tabanlı bir Derin Sistem Modeli (KSA-DSM) önerilmiştir. Önerilen bu modelde klasik KSA katmanlarına ek olarak bozulmuş görüntü verilerini katmanlara süren bir artırma katmanı tasarlanmıştır. Çalışmanın performans testleri için gerçek dünyadan elde edilen çeşitli yüzey görüntüleri CURET veri tabanından seçilerek kullanılmıştır. Elde edilen sonuçlar sunularak tartışılmıştır.

**Anahtar Kelimeler**—derin öğrenme; konvolüsyonel sinir ağı; gerçek dünya doku görüntüleri.

**Abstract**—Images obtained from the real world environments usually have various distortions in image quality. For example, when an object in motion is filmed, or when an environment is being filmed on the move, motion tracking effects occur on the image. Increasing the recognition performance of expert systems, which perform image recognition on data obtained under such conditions, is an important research area. In this study, we propose a Convolutional Neural Network (CNN) based Deep System Model (CNN-DSM) for accurate classification of images under challenging conditions. In the proposed model, a new layer is designed in addition to the classical CNN layers. This layer works as an enhancement layer. For the performance evaluations, various real world surface images were selected from the CURET database. Finally, results are presented and discussed.

**Keywords**—deep learning; convolutional neural networks; real-world texture images.

### I. GİRİŞ

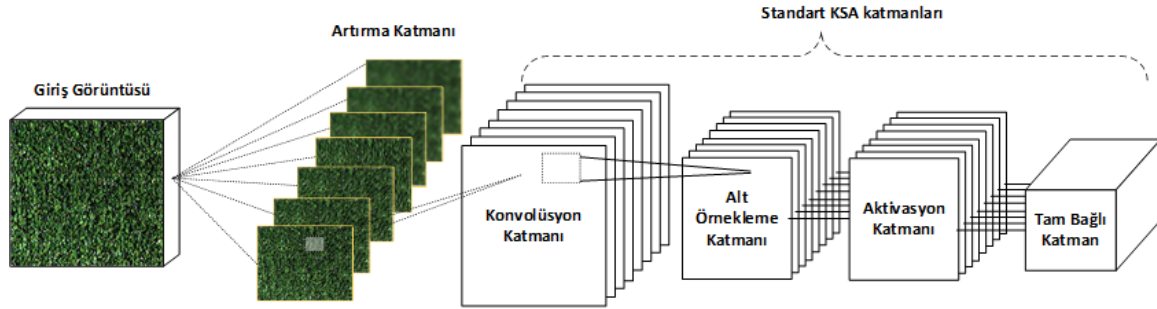
Düşük kaliteli görüntülerdeki gerçek dünya doku bilgilerinin algılanması ve tanınması, görüntü işleme ve bilgisayar görmesinde önemli bir yere sahiptir. Özellikle; sürücüsüz araçlar, robotik, insan-robot etkileşimi ve insan-makine sistemleri gibi birçok uygulamada çok önemli bir çalışma alanıdır [1, 2]. Gerçek dünya yüzeylerin doku analizi ile ilgili ilk çalışmalar, Dana vd. tarafından sınırlı ortamda oluşturulan CURET veri seti ile yapılmıştır [3]. Bu veri kümesi 2009 yılında, Varma ve Zisserman tarafından doku sınıflandırma için kullanılmıştır [4]. 2010'da Sharan vd. dokulara ilişkin farklı özellikler kullanarak bir material veri kümesi üzerinde doku sınıflandırma işlemi gerçekleştirmişlerdir, fakat düşük bir başarımla elde etmişlerdir [5].

Son zamanlarda doku sınıflandırma görevinde de, derin öğrenme mimarileri sıklıkla kullanılmaya başlanmıştır. Titive vd. [6] sınıf sayısı az olan düşük çözünürlüklü görüntüler içeren Brodatz doku veri kümesine KSA'ları uygulamıştır. Daha derin bir çalışma, 2014'te Cimpoli vd. tarafından yapılmıştır. Cimpoli vd. FV-CNN isimli bir yöntem önermiştir [7]. Doğrudan KSA'ları kullanmak yerine Fisher Vectors (FV) yöntemini kullanmıştır. FV-CNN'nin sınıflandırma bölümünde, doku veri kümesini doğrudan öğrenmek için KSA'lar kullanılmıştır.

Doku analiz ve sınıflandırma işlemlerinde en önemli problemlerden birisi görüntü üzerinde meydana gelen hareket-izi veya sis etkisidir. Gerçek hayattaki görüntülerde hareket-izi sorunu genellikle kameranın sallanması [8, 9] veya nesnenin hareketi [10, 11] gibi durumların kombinasyonu ile ortaya çıkar. Bu etkiler daha fazla maliyet bedelinde kaliteli görüntü sensörü ile veya pozlama süresini ayarlayarak azaltılabilir. Benzer şekilde sis etkisi de görüntünün elde edildiği ortam veya kamera üzerinde hava şartlarından oluşan buğulanma gibi durumlarda ortaya çıkabilmektedir.

Bu çalışmada, hareket-izi ve sis etkisinin baskın olduğu gerçek dünya yüzey dokularında, dayanıklı ve yüksek





Şekil 2. Önerilen KSA-DSM modeline ait blok gösterim.

Giriş görüntüleri öncelikle artırma katmanına verilmekte ve bu katmanda görüntü çeşitli çekirdek parametreleri ile konvolüsyon işlemine tabi tutulmaktadır. Bu çekirdek parametreleri sabittir ve görüntü üzerinde zorlu koşul oluşturacak biçimde belirlenmiştir. Böylece bu katman çıkışında görüntülere ait belirli oranlarda/açılarda hareket-izi ve sis etkileri içeren alt örnekler elde edilmektedir.

#### A. CURET Veri Seti:

Önerilen tanıma sisteminin performans değerlendirmeleri için gerçek-dünya dokuları içeren Curet [23] veri seti kullanılmıştır. Bu veri seti toplam 61 farklı gerçek-dünya yüzeylerine ait görüntüler içermektedir. Farklı kategorilerde veri barındıran geniş bir koleksiyona sahiptir. Bu çalışmada, 61 veri sınıfı içerisinde çalışmanın amacına yönelik olarak 22 adet yüzey görüntüsü seçilerek kullanılmıştır. Şekil 3'te çalışmada kullanılan Curet veri seti içerisindeki yüzey görüntüleri verilmiştir.



Şekil 3. Curet veri seti içerisindeki yüzey doku görüntüleri.

#### B. Deneysel Sonuçlar:

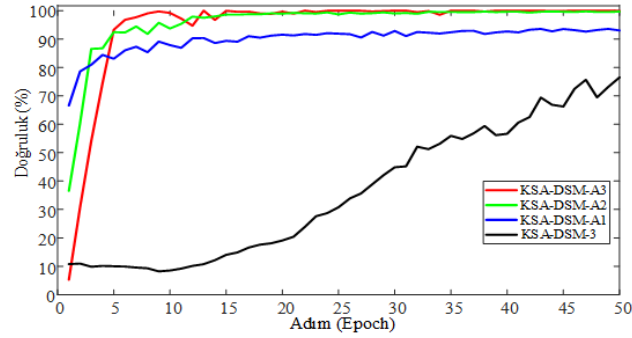
Curet veri seti üzerindeki sınıflandırılma çalışmaları için dört farklı KSA-DSM modeli hazırlanmıştır. Bu modeller içerdikleri alt-örnekleme katman sayılarına göre; KSA-DSM-A1, KSA-DSM-A2 ve KSA-DSM-A3 şeklinde isimlendirilmişlerdir. Ayrıca KSA-DSM-A3 modelinden artırma katmanını çıkarılmış olan KSA-DSM-3 modeli artırma katmanının etkisini ortaya koymak için hazırlanmıştır. Bu modeller ve içerdikleri parametreler Şekil 4'te detaylı olarak verilmiştir.

KSA-DSM-A1	KSA-DSM-A2	KSA-DSM-3	KSA-DSM-A3
Augment 64,64,3 @48	Augment 64,64,3 @48	Conv 8,8,3 @32	Augment 64,64,3 @48
Conv 8,8,3 @32	Conv 8,8,3 @32	Max Pool 5,5	Conv 8,8,3 @32
Max Pool 5,5	Max Pool 5,5	ReLU	Max Pool 5,5
ReLU	ReLU	Conv 3,3,32 @32	ReLU
Conv 1,1,32 @22	Conv 3,3,32 @32	Avg Pool 3,3	Conv 3,3,32 @32
SoftMax	Avg Pool 3,3	Conv 5,5,32 @32	Avg Pool 3,3
	Conv 1,1,32 @22	Avg Pool 3,3	Conv 5,5,32 @32
	SoftMax	Conv 1,1,32 @22	Avg Pool 3,3
		SoftMax	Conv 1,1,32 @22
			SoftMax

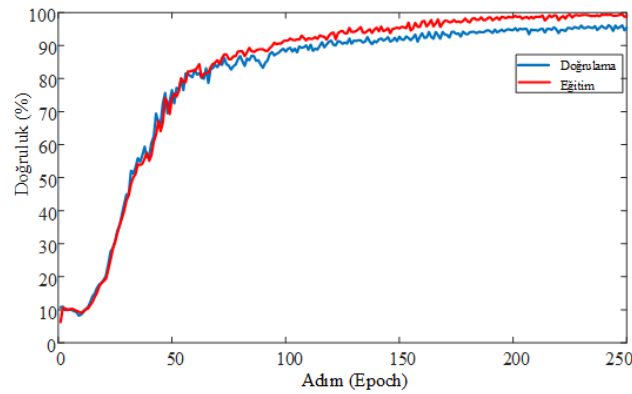
Şekil 4. Curet veri seti için hazırlanan KSA-DSM modelleri ve parametreleri.

Bu modellerin bir birlerine göre performans karşılaştırmaları yapılmış ve veri seti üzerinde elde ettikleri başarımlar tespit edilmiştir. Bu amaçla Curet veri seti üzerinde gerçekleştirilen tanıma çalışmaları için 22 yüzey görüntüsüne ait örnekler kullanılmıştır. Her sınıf için 64x64x3 boyutlarında 100 adet alt görüntü üretilmiştir. Metodların eğitimi aşamasında 550 adet (22 sınıf x 25 örnek) ve test aşamasında 1650 (22 sınıf x 75 örnek) görüntü kullanılmıştır. Bunun yanı sıra her sınıfa ait 25 eğitim

verisinin %80'i eğitim %20'si doğrulama verisi olarak değerlendirilmiştir. KSA-DSM modellerinin öğrenme süreçlerine ait performans grafikleri Şekil 5'de gösterilmiştir.



(a) Tüm modellerin öğrenme performansları.



(b) KSA-DSM-3 modeline ait eğitim ve doğrulama performansı.

Şekil 5. KSA-DSM modellerinin veri seti üzerindeki öğrenme performansları.

Şekil 5'de, görüldüğü gibi KSA-DSM-3 modelinin eğitimi yaklaşık 200 adımda tamamlanırken doğrulama aşaması tam olarak başarıya ulaşamamıştır. Bunun yanında, artırma katmanı eklenmiş olan KSA-DSM-A3 modeli 15 adım gibi kısa bir sürede hem eğitim hem de doğrulama aşamalarını başarıyla tamamlamıştır. Bu öğrenme grafiklerinden görüleceği gibi çalışmada önerilen artırma katmanı öğrenme sürecini hem başarı hem de hız açısından oldukça iyileştirmiştir.

Elde edilen performans grafiğine göre bu veri seti için en uygun KSA-DSM yapısının KSA-DSM-A3 modeli olduğu söylenebilir. Bu nedenle sonraki performans testlerinde bu modelin artırma katmanlı olan KSA-DSM-A3 ve artırma katmanı olmayan KSA-DSM-3 modelleri kullanılmıştır. Her iki modelin hareket-izi ve sis etkisi altındaki tanıma başarımlarını değerlendirmek için test verilerine farklı oran ve açılarla hem sis etkisi hem de hareket-izi etkisi eklenmiştir. Tablo 1'de her iki KSA-DSM modelinin bu test veri setleri üzerindeki tanıma başarımları verilmiştir.

Tablo 1. Hareket-izi eklenmiş test veri seti için modellerin tanıma performansları.

Model	Açı ( $\alpha$ )	Uzaklık				
		4	8	16	32	48
KSA-DSM-3	30°	%83.4	%57.45	%42.30	%35.87	%35.75
	45°	%86.06	%65.33	%48.24	%38.66	%36.18
	60°	%88.96	%79.75	%58.90	%44.00	%38.66
	90°	%94.96	%89.39	%78.48	%66.00	%55.87
KSA-DSM-A3	30°	%99.33	%99.33	%98.06	%92.72	%85.93
	45°	%99.27	%99.51	%99.09	%93.39	%86.48
	60°	%99.09	%99.51	%99.09	%93.45	%86.96
	90°	%99.09	%99.21	%98.90	%96.42	%91.15

Tablo 1'de elde edilen sonuçlar, KSA-DSM-A3 modelinin hareket-izi etkisi altında görüntüleri oldukça başarılı bir şekilde tanıdığını göstermektedir. KSA-DSM-3 modeli ise uzaklık değerinin 4 olduğu durumda başarımlarını göstermesine rağmen uzaklık değeri arttıkça verileri tanımda oldukça başarısız olduğu görülmektedir. Örneğin 30°'lik hareket-izi oluşan verilerde uzaklık 4 iken %83.45 başarımlar sağlarken uzaklık değeri 48 olduğunda bu oran %35.75'e kadar düşmüştür. KSA-DSM-A3 modelinde ise bu oran %99.33 değerinden sadece %85.93 değerine düşmüştür. Bu sonuç önerilen artırma katmanlı KSA-DSM-A3 yapısının hareket-izine dayanıklılığını göstermektedir. Cüret veri seti üzerinde yapılan diğer bir performans değerlendirmesi ise sis etkisi eklenmiş olan test verileri üzerinde gerçekleştirilmiştir. Sis etkisi için  $\sigma=4, 8, 16, 32$  ve 48 oranları kullanılmıştır. Tablo 2'de test verileri için KSA modellerinin performansları gösterilmiştir.

Tablo 2. Cüret veri seti için sis etkisi eklenmiş test görüntüleri üzerinde KSA-DSM modellerinin başarımları.

Orjinal Görüntü	$\sigma=4$	$\sigma=8$	$\sigma=16$	$\sigma=32$	$\sigma=48$
Model	Sınıflama Doğruluğu (%)				
KSA-DSM-A3	99.09	95.93	81.33	64.24	54.06
KSA-DSM-3	38.24	30.84	29.87	28.54	28.36

Tablo 2’de elde edilen sonuçlar, KSA-DSM-3 modelinin hafif bir sis etkisi altında bile görüntüleri tanıyamadığını buna karşın artırma katmanı içeren KSA-DSM-A3 modelinin  $\sigma=16$  değerine kadar %80 üstünde bir başarımla gösterdiğini ortaya koymaktadır.  $\sigma=32$  ve üstü değerlerde görüntülerin üzerini oldukça kalın bir sis kapladığı için görüntüler artık belirsizleşmiştir. Buna rağmen artırma katmanı eklenmiş olan KSA-DSM-A3 modeli %50 üzerinde bir oranla görüntüleri tanıyabilmektedir.

#### IV. SONUÇLAR

Bu çalışmada, gerçek dünya doku görüntülerinin zorlu koşullar altında tanınması için KSA-DSM isimli bir tanıma yöntemi geliştirilmiştir. KSA tabanlı olan bu yöntem kapsamında yeni bir artırma katmanı tasarlanmıştır. Bu artırma katmanı sayesinde tanıma performansının iyileştirilmesi sağlanmıştır. Çalışmanın temel amacı, hareket-izi ve sis etkisi gibi zorlu koşullar altında görüntülerin yüksek doğrulukta tanınmasıdır. DSM-KSA yöntemi içerisinde dört farklı model hazırlanmış ve bu modellerin Curret veri seti üzerinde performans değerlendirmeleri gerçekleştirilmiştir. Elde edilen sonuçlar değerlendirildiğinde, önerilen artırma katmanlı KSA-DSM modelin hem normal verilerdeki başarımlı artırdığı hem de zorlu koşullar altında bile görüntüleri yüksek doğrulukta sınıflandırdığı gözlemlenmiştir.

#### YAZAR KATKILARI

*Özal Yıldırım* çalışmanın sorumlu yazarı olup yöntem ve deneysel çalışma kısımlarında etkin olarak katkı sunmuştur. *Ayşegül Uçar*, makale konusunun literatürdeki yeri ile ilgili kapsamlı çalışmalar gerçekleştirerek yöntem konusunda ve giriş kısmında çalışmalar gerçekleştirmiştir. *Ulaş Baran Baloğlu* ise makale konusunun arkaplanı ve deneysel çalışmalarda katkılar sunmuştur.

#### KAYNAKÇA

- [1] P. Azad, “Visual perception for manipulation and imitation in humanoid robots”, Springer Science & Business Media, 4, 2009.
- [2] R. Manduchi, A.Castano, A. Talukder, L. Matthies, “Obstacle detection and terrain classification for autonomous off-road navigation”, Autonomous robots, 18(1), 81-102, 2005.
- [3] K.J. Dana, B. Van Ginneken, S.K. Nayar, J.J. Koenderink, “Reflectance and texture of real-world surfaces”, ACM Transactions On Graphics (TOG), 18(1), 1-34, 1999.
- [4] M. Varma, A. Zisserman, “A statistical approach to material classification using image patch exemplars”, IEEE transactions on pattern analysis and machine intelligence, 31(11), 2032-2047, 2009.
- [5] L. Sharan, R. Rosenholtz, E. Adelson, “Material perception: What can you see in a brief glance?”, Journal of Vision, 9(8), 784-784, 2009.
- [6] F. H. C. Tivive and A. Bouzerdoum, “Texture classification using convolutional neural networks,” in TENCON 2006, IEEE Region 10 Conference, 2006, pp. 1–4.
- [7] M. Cimpoi, S. Maji, I. Kokkinos, A. Vedaldi, “Deep filter banks for texture recognition, description, and segmentation”, International Journal of Computer Vision, 118(1), 65-94, 2016.
- [8] O. Whyte, J. Sivic, A. Zisserman, J. Ponce, “Non-uniform deblurring for shaken images”, International journal of computer vision, 98(2), 168-186, 2012.
- [9] S. Zheng, L. Xu, J. Jia, “Forward motion deblurring”, In Proceedings of the IEEE international conference on computer vision, pp. 1465-1472, 2013.
- [10] T. Hyun Kim, B. Ahn, K. Mu Lee, “Dynamic scene deblurring”, In Proceedings of the IEEE International Conference on Computer Vision, pp. 3160-3167, 2013.
- [11] J. Pan, Z. Hu, Z. Su, H.Y. Lee, M.H. Yang, “Soft-segmentation guided object motion deblurring”, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 459-468, 2016.
- [12] Y. Bengio, “Learning deep architectures for AI. Foundations and trends in Machine Learning”, 2(1), 1-127, 2009.
- [13] A. Krizhevsky, I. Sutskever, G.E. Hinton, “Imagenet classification with deep convolutional neural networks”, In Advances in neural information processing systems, 1097-1105, 2012.
- [14] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, “Gradient-based learning applied to document recognition”, Proceedings of the IEEE, 86(11), 2278-2324, 1998.
- [15] S. Lawrence, C.L. Giles, A.C. Tsoi, A.D. Back, “Face recognition: A convolutional neural-network approach”, IEEE transactions on neural networks, 8(1), 98-113, 1997.
- [16] B. Fasel, “Robust face analysis using convolutional neural networks”, IEEE Proceedings. 16th International Conference on Pattern Recognition, 2002.
- [17] S. Ji, “3D convolutional neural networks for human action recognition”, IEEE transactions on pattern analysis and machine intelligence, 35(1), 221-231, 2013.
- [18] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, L. Fei-Fei, “Large-scale video classification with convolutional neural networks”, In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 1725-1732, 2014.
- [19] K. Jarrett, K. Kavukcuoglu, M. Ranzato, and Y. LeCun., “What is the best multi-stage architecture for object recognition?”, In Proc. International Conference on Computer Vision (ICCV’09), 2009.
- [20] R. Girshick, J. Donahue, T. Darrell, J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation”, In Proceedings of the IEEE conference on computer vision and pattern recognition, 580-587, 2014.
- [21] O. Abdel-Hamid, L. Deng, D. Yu, “Exploring convolutional neural network structures and optimization techniques for speech recognition”, In Interspeech, 3366-3370, 2013.
- [22] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, ... & S. Dieleman, “Mastering the game of Go with deep neural networks and tree search”, Nature, 529(7587), 484-489, 2016.