

# Gri Seviye Birliktelik Matrisi Öznitelikleri ve Manifold Öğrenme Yardımıyla Histoloji Görüntülerinde Otomatik Doku Sınıflandırılması

## Automated Clustering Of Histology Slide Texture Using Co-Occurrence Based Grayscale Image Features and Manifold Learning

Devrim Önder, Bilge Karaçalı

Elektrik ve Elektronik Mühendisliği Bölümü,  
İzmir Yüksek Teknoloji Enstitüsü  
{devrimonder, bilgekaracali}@iyte.edu.tr

### Özetçe

*Bu çalışmanın amacı, histoloji yansılarında elde edilmiş gri seviyeli görüntülerde manifold öğrenme yöntemi kullanarak otomatik doku sınıflandırmasını gerçekleştirmektir.*

*Dokulara ait öznitelik vektörleri, lokal birliktelik matrisleri kullanılarak elde edilmiş ve öznitelik vektör uzayı boyutları, İsomap boyut düşürme yöntemi ile azaltılmıştır. Elde edilen düşük boyutlu uzayda, k-ortalama kümeleme yöntemi ile birbirinden farklı doku kümeleri belirlenmiştir.*

*Bu çalışmada, insan böbreğine ait histoloji yansuları kullanılmıştır. Elde edilen boyutu düşürülmüş öznitelik vektörleri ile doku kümeleri sonuç olarak verilmiştir. Terimler — Doku, birliktelik, Isomap, boyut düşürme.*

### Abstract

*The aim of this work is to perform automated texture classification of histology slides using grayscale images and manifold learning method.*

*Texture feature vectors were obtained using local gray scale co-occurrence matrices and the dimension of the feature vector space was lowered using Isomap dimension reduction. In a lower dimension feature space, k-means clustering operation was performed in order to provide separate texture clusters.*

*In this work, experimental results were obtained using human kidney histology slides. Corresponding feature vectors and determined texture types were given as results.*

*Index Terms — Texture, co-occurrence, Isomap, dimensionality reduction.*

### 1. Giriş

Histoloji yansuları içinde doku sınıflandırılması yapmak amacıyla çeşitli yöntemler üzerinde çalışılmıştır [1] ve [2].

Bu çalışmada, hücre yansuları içerisindeki farklı doku yapıları, ilgili dijital görüntüleri ait doku analizi kaynaklı gri-seviye birliktelik matrislerinden elde edilen Haralick öznitelik vektörleri kullanılarak sınıflandırılmıştır [3] ve [4]. Hesaplanan çok boyutlu öznitelik vektörleri, İsomap boyut düşürme yöntemi ile daha düşük boyutlu bir uzayda ifade edilmiştir. Bu düşük boyutlu uzayda uygulanan k-ortalama kümeleme yöntemi ile her bir vektörün ait olduğu doku kümeleri belirlenmiştir.

Önerilen yöntem, bir sonraki bölümde detaylandırılmıştır. Bölüm 3’de, yöntemin, gerçek böbrek histoloji görüntülerine uygulanmasıyla elde edilen sonuçlar sunulmaktadır. 4. Bölüm’de ise, elde edilen sonuçlar ve gelecekte yürütülecek çalışmalar anlatılmaktadır.

## 2. Yaklaşım

### 2.1. Histoloji görüntüleri

Yapılan deneylerde, Iowa Üniversitesi sitesinden (<http://www.path.uiowa.edu/virtualslidebox>) alınmış, beş farklı insan böbreği histoloji görüntüsü kullanılmıştır.

### 2.2. Doku öznitelik vektörlerinin oluşturulması

Doku analizi, görüntü işleme alanındaki temel çalışma konularından birisidir. Her bir farklı doku, birbirine benzer ton ve yapıdaki bitişik görüntü noktalarından oluşur.

Bir görüntü bölgesi, içerisindeki lokal ton özellikleri sabit, yavaş değişiyor veya periyodik ise, belli bir dokuya ait olarak değerlendirilir. Doku sınıflandırma, bir görüntüyü, dokulara ait parçalara ayırmayı amaçlar. Bu sayede, her bir doku ile ilişkili olan nesnelere elde edilebilir.

Doku tanımlama yaklaşımları aşağıdaki şekilde listelenebilir:

- İstatistiksel: Her bir doku, bir öznitelik vektörü olarak tanımlanır.

- Yapısal: Doku elemanları, belli bir anlam taşıyan bölgeler olarak tanımlanır.
- Sözdizimsel: Doku elemanları belli bir dilbilgisi kullanılarak ilişkilendirilir.
- Melez: İstatistiksel ve sözdizimsel yaklaşımlar bir arada kullanılır.

Birliktelik matrisi, öznelikler arasındaki ilişkileri belirleyen ve en sık olarak kullanılan doku tanımlayıcısıdır. Birliktelik matrislerinin özellikleri şöyle listelenebilir;

1. Doku tanımlamasına çok boyutta olanak tanımak,
2. Dokular arasındaki farklılığın doğrudan istatistiksel ölçüm yöntemleri ile elde edilmesini sağlamak,
3. Haralick öznelikleri gibi ikincil özneliklerin elde edilmesini sağlamak.

Haralick doku öznelikleri en çok kullanılan birliktelik matrisi çıktısıdır. Haralick doku öznelikleri; açılal ikincil moment, ters ikincil moment, zıtlık, varyans, yerel varyans, toplam ortalama, toplam varyans, toplam varyans, düzensizlik, fark varyansı, fark düzensizliği, maksimum olasılık ve ilinti olarak verilebilir.

Yapılan deneylerde 128x128 veya 256x256 gibi farklı nokta boyutunda kare bloklar kullanılarak birliktelik matrisleri hesaplanmıştır. Bu hesaplamada, gri seviyede 16'lık nicemleme yapılmıştır. Sözkonusu birliktelik matris elemanları kullanılarak, formülleri Tablo 1'de verilen beş adet Haralick özneliği elde edilmiştir [5].  $p(i,j)$  her bir birliktelik matris elemanına karşılık gelmektedir.

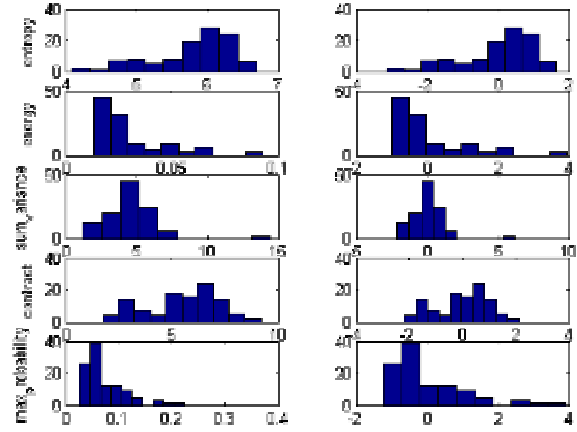
**Tablo 1. Doku özneliklerinin hesaplanması**

düzensizlik	$-\sum_i \sum_j p(i,j) \log(p(i,j))$
enerji	$\sum_i \sum_j p(i,j)^2$
toplam varyans	$\sum_i \sum_j (i - \mu)^2 p(i,j),$ $\mu = \sum_i \sum_j p(i,j)$
zıtlık	$\sum_n n^2 \sum_i \sum_j p(i,j),  i-j =n$
maksimum olasılık	$\max(p(i,j))$

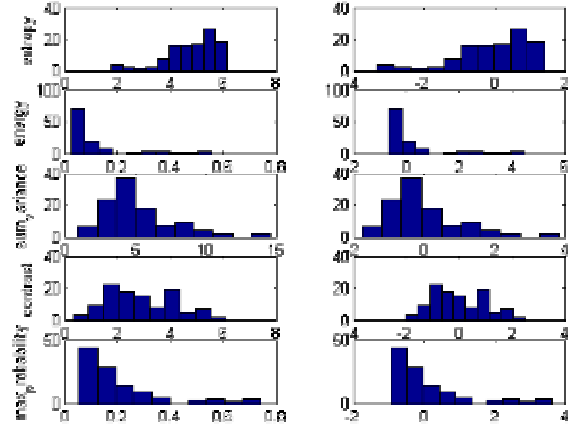
Görüntülerden elde edilen öznelik vektörlerine düzgünleştirme işlemi uygulanmıştır. Düzgünleştirme amacıyla vektör bileşenlerin ortalama ve varyans değerleri kullanılmıştır. Her bir vektör bileşeninin ortalama değeri, sözkonusu bileşen değerlerinden çıkarılarak 0.0 ortalamaya sahip bileşenler elde edilmiştir. Bunu takiben, her bir vektör bileşeninin

varyans değeri hesaplanmış ve ilgili tüm vektör bileşenleri bu değere bölünerek varyansı 1.0 olan bileşenler elde edilmiştir.

Örnek olarak, bir vektör düzgünleştirme işleminin sonuçları Şekil 1'de verilmiştir. İki farklı test görüntüsünden çıkarılan özgün ve düzgünleştirilmiş vektör bileşenleri dağılım eğrileri Şekil 1.a'da ve 1.b'de verilmiştir. Şekillerdeki ilk kolon özgün, ikinci kolon düzgünleştirilmiş değerlere ait grafiklerdir.



(a)



(b)

**Şekil 1. Vektör bileşenlerinin düzgünleştirme öncesi ve sonrası dağılım eğrileri.**

İkinci kolonlardaki grafiklerin sıfır ortalamaya sahip olduğuna dikkat edilmelidir. Ayrıca, her iki vektör kümesindeki özgün enerji bileşenlerinin (ikinci satır grafikleri) farklı aralıklarda değerler aldığı görülürken, düzgünleştirilme sonrası aynı değer aralığı içerisine kaldığı görülebilir.

### 2.3 İsomap yöntemi ile doğrusal olmayan boyut düşürme

Dokular, Haralick öznitelikleri kullanılarak yüksek boyutlu bir uzayda vektörler olarak ifade edilebilir. Yüksek boyutlu, iki veya üç boyuttan fazlasına ihtiyaç duyan bir bilginin yorumlanması karmaşık olmaktadır. Bu karmaşıklığı düşürme yöntemlerinde birisi, bilginin yüksek boyutlu gösterim içinde, doğrusal olmayan bir manifold üzerinde yer aldığını varsaymaktadır. Eğer manifold, göreceli olarak daha düşük boyuta indirgenbilirse, elde edilen bilgi daha düşük boyutlu bir uzayda izlenebilecektir. Bu sebeple, çok boyutlu vektör bilginin gözlenmesi şu anda oldukça cazip Makine Öğrenimi konularındandır. Mevcut doğrusal analiz sistemleri, doğrusal olmayan manifoldlar üzerindeki yerel değişimleri göz ardı etmekte ve bu belli bir kayıp meydana getirmektedir. Yüksek başarılı tanıma sistemleri için sözkonusu sorunun giderilmesi gerekmektedir.

Son yıllarda doğrusal olmayan manifoldlar konusunda çalışmalar hızlanmış ve bu alanda mevcut kaynak sayısı artmıştır. Aşağıda manifold öğrenme ve doğrusal olmayan boyut düşürme konusundaki önemli yöntemler listelenmiştir [6]:

1. Doğrusal yöntemler;
  - Bağımsız bileşen analizi,
  - Temel bileşen analizi (Karhunen-Loève, KL dönüşümü olarak da bilinir.),
  - Tekil değer çözümlenmesi,
  - Faktör analizi.
2. Doğrusal olmayan yöntemler;
  - Çekirdek temel bileşen analizi [7]
  - Lokal doğrusal yerleştirme [8],
  - Lokal teğet uzayı öteleme [9].
3. Yakınlık matrislerine dayalı yöntemler.

Yukarıda verilen yöntemler, yakınlık bilgisinin hesaplanmasındaki yaklaşımlarından dolayı farklılık göstermektedir. Yakınlık bilgisinin hesaplanması için kullanılan yöntemler olarak İsomap, lokal doğrusal yerleştirme, maksimum varyans açılımını sayabiliriz [10].

İsomap yöntemi, ağırlıklandırılmış grafikler üzerindeki kesel uzaklıkların, klasik çok boyutlu ölçekler ile ilişkilendirildiği, sıklıkla kullanılan düşük boyutlu yerleştirme yöntemlerinden birisidir. Bu yöntem, çok boyutlu bir vektör kümesinin düşük boyutlu, yakın eş ölçekli yerleştirmesini hesaplamakta kullanılır.

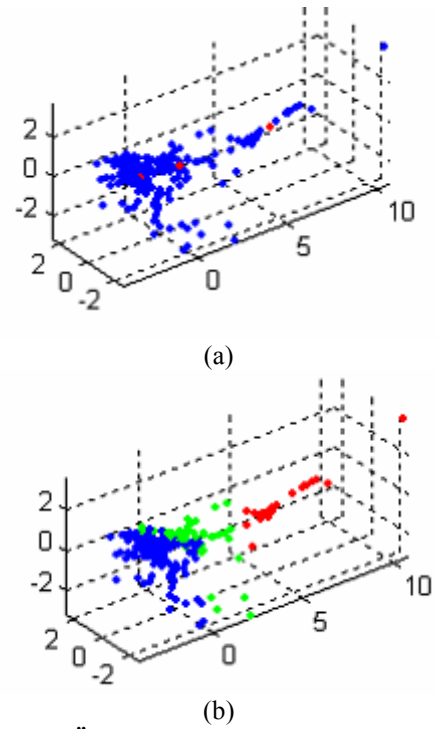
İsomap, ağırlıklandırılmış bir grafik üzerinde Öklid uzaklıkları yerine Dijkstra'nın en kısa kesel uzaklıklarını kullanan bir eş ölçekli eşleştirme yapmaktadır. Yöntem, her bir veri noktasının manifold üzerindeki tahmini komşu noktasını kullanarak sözkonusu manifoldun iç geometrisini tahmin eden karmaşık olmayan bir yöntemdir. Gözlem uzayında ve iç yerleştirme uzayındaki eş ölçek değerlerinin korunması temeline dayanır. Farklı boyutlarda ve yapıdaki bir çok veri yığnında etkin olarak kullanılabilir.

### 3. Sonuçlar

Yapılan deneylerde, elde edilen öznitelik vektörleri, İsomap manifold öğrenme yöntemi ile beş boyutlu uzaydan üç boyutlu uzaya düşürülmüştür [11]. Sözkonusu düşük boyutlu vektör uzayında, her bir vektörün ait olduğu farklı kümeler k-ortalama kümeleme yöntemi ile belirlenmiştir.

Bir görüntüye ait, üç boyutlu uzaya düşürülmüş doku öznitelik vektörleri, Şekil 2.a'da verilmiştir. Kümeleme işlemi sonrasında, her bir vektör, ait olduğu kümenin renginde Şekil 2.b'de gösterilmiştir.

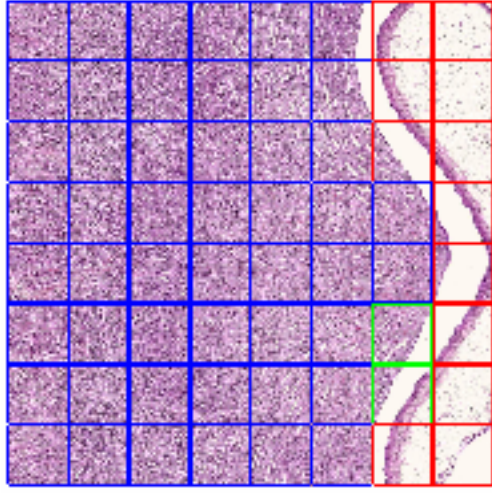
Her bir analiz bloğu için belirlenen doku sınıfları, çalışılan görüntülerden ikisi üzerinde farklı renklerde, Şekil 3'de gösterilmiştir.



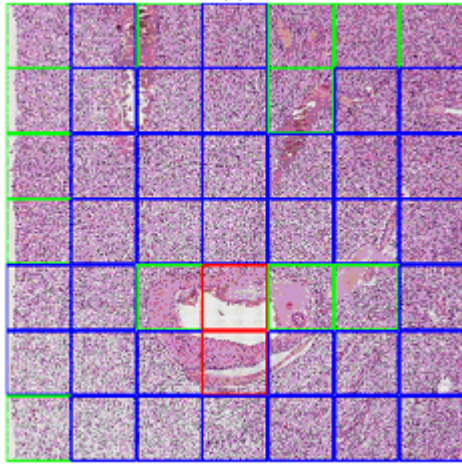
Şekil 2. Üç boyutlu öznitelik vektörleri ve kümeleme

### 4. Tartışma

Yapılan çalışmada, histoloji görüntülerinden, gri seviye birliktelik matrisleri hesaplanmış ve seçilen beş Haralick özniteliği kullanılarak öznitelik vektörleri elde edilmiştir. İsomap manifold öğrenme yöntemi kullanılarak, öznitelik uzayı boyutu düşürülmüş ve k-ortalama yöntemi ile kümeleme yapılmıştır. Her bir kümeye karşılık gelen histoloji görüntü bölgeleri incelenmiş ve yöntemin dokuların sınıflandırmasında başarılı olduğu görülmüştür.



(a)



(b)

**Şekil 3. Histoloji görüntülerinde doku sınıfları**

İsomap manifold öğrenme yöntemi, elde edilen öznelik vektörlerini daha düşük boyutlarda yorumlama olanağı vermekte ve böylelikle, sınıflandırma işlemlerinin etkin biçimde yapılmasını sağlamaktadır.

Yapılan çalışmada, serbest kaynaklardan alınan histoloji görüntüleri kullanılmış ve gri-seviyede düzgünleştirme yapılmamıştır. Çalışmanın, renk dağılımı ve resimleme parametreleri kontrol altında alınmış başka görüntü kütüphanelerinde tekrarlanması başarıyı arttıracaktır.

Yapılan çalışmada kullanılan öznelik sayısı artırılabilir olup, diğer Haralick öznelikleri veya gri-seviye bilgisine ek olarak, renk bilgisi de kullanılabilir. Kullanılan yöntem, bu açıdan genişletilebilir özelliktedir. İlerideki çalışmalarda, doku sınıflandırma için gradyan görüntünün kullanılması da düşünülebilir. Bu durumda, herhangi bir gri-seviye düzgünleştirme işlemine gerek olmayacağı varsayılmaktadır. Çalışmanın devamında, sağlıklı görüntülerle birlikte, kanserli histoloji görüntülerinin kullanılarak, anormal görüntü

bölgelerinin, ayrı doku sınıfları şeklinde belirlenmesi hedeflenmektedir.

## Kaynakça

- [1] Automated detection of regions of interest for tissue microarray experiments: an image texture analysis, Bilge Karaçali and Aydın Tözeren, BMC Medical Imaging 2007, 9 March 2007, doi:10.1186/1471-2342-7-2.
- [2] Dreyer T, Doudkine A, Popella C: Nuclear shape and nuclear texture analysis in normal and neoplastic laryngeal epithelium. The First South-East European Symposium on Interdisciplinary Approaches in Fractal Analysis 2003:241-244.
- [3] Robert M. Haralick, K Shanmugam and Its'Hak Dinstein(1979). "Textural Features for Image Classification.", IEEE Transactions on Systems, Man, and Cybernetics.
- [4] Haralick, R.M., (1979), "Statistical and Structural Approaches to Texture", Proceedings of the IEEE, Vol. 67, No. 5, s. 786-804.
- [5] Eizan Miyamoto and Thomas Merryman Jr., "Fast Calculation Of Haralick Texture Features", Human Computer Interaction Institute, Dept. of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213.
- [6] Jose A. Costa and Alfred O. Hero III, Manifold Learning Using Euclidean K -Nearest Neighbor Graphs, Dept. of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109.
- [7] B. Schölkopf, A. Smola, K.-R. Muller, Kernel Principal Component Analysis, In: Bernhard Schölkopf, Christopher J. C. Burges, Alexander J. Smola (Eds.), Advances in Kernel Methods-Support Vector Learning, 1999, MIT Press Cambridge, MA, USA, 327-352. ISBN 0-262-19416-3.
- [8] S. T. Roweis, and K. S. Lawrence, (2000), "Nonlinear Dimensionality Reduction by locally linear embedding," Science, 290, pp. 2323-2326.
- [9] Zhenyue Zhang and Hongyuan Zha, Principal Manifolds and Nonlinear Dimension Reduction via Local Tangent Space Alignment, SIAM Journal on Scientific Computing 26 (1) (2005), 313 - 338.
- [10] V. de Silva, J. C. Langford, A Global Geometric Framework for Nonlinear Dimensionality Reduction, Science 290, (2000), 2319-2323.
- [11] A Global Geometric Framework for Nonlinear Dimensionality Reduction, J. B. Tenenbaum, V. de Silva and J. C. Langford, Science 290 (5500): 2319-2323, 22 December 2000, <http://web.mit.edu/cocosci/isomap/isomap.html>.