



Pulmoner Emboli Veri Setinde Özellik İndirgeme Feature Selection in Pulmoner Emboli Dataset

Haydar Özkan¹, Gökalp Tulum² Onur Osman³, Sinan Şahin⁴

¹ Biyomedikal Mühendisliği
Fatih Sultan Mehmet Vakıf Üniversitesi
haydarozkan79@gmail.com

² Elektrik Elektronik Mühendisliği
Yeni Yüzyıl Üniversitesi
gokalptulum@gmail.com

³ Elektrik Elektronik Mühendisliği
İstanbul Arel Üniversitesi
onur.osman@gmail.com

⁴ Radyoloji Anabilim Dalı
Dr. Siyami Ersek Göğüs Kalp ve Damar Cerrahisi Eğitim ve Araştırma Hastanesi
sinan.sahin@gediknet.com

Özetçe

Bu çalışmada Bilgisayarlı Tomografi Anjiyografi (BTA) görüntülerinde Pulmoner Emboli (PE) adayları için öncelikle özellik çıkartımı gerçekleştirilmiştir. Elde edilen özellik kümesine Temel Bileşen Analizi (TBA) uygulanarak özellik seçimi uygulanmıştır. Elde edilen özellik kümesi ve TBA sonrası elde edilen indirgenmiş küme Yapay Sinir Ağı (YSA) ile sınıflandırılarak PE'lerin tespiti yapılmıştır. Çalışmada toplamda 450 PE'si olan 33 hastaya ait veri seti kullanılmıştır.

Abstract

In this paper a feature extraction matrix is extracted from Pulmonary Embolism candidates in Computed Tomography Angiography (CTA) images. Also Principal Component Analysis (PCA) is used as a feature selection method. For a given feature extraction and selection matrices, Artificial Neural Network (ANN) is used as a classifier. In this work, 33 patients with 450 PE's are considered.

1. Giriş

PE, akciğerlere kan götüren pulmoner arterlerin tıkanmasına bağlı olarak ortaya çıkan tanısı güç bir hastalık türüdür. Akut pulmoner emboli olarak tanımlanan ve bir kaç saat gibi kısa bir sürede gelişen türünde erken tanı ve tedavinin hayat kurtarıcı olduğu görülmektedir [1]. Günümüzde, Bilgisayar-Destekli Tespit (CAD) sistemleri

BTA görüntüleri üzerinde tanı ve teşhis koymada radyologlara yardımcı olacak sistemler olarak kullanılmaktadır. CAD ile teşhis sürelerinin düşürülmesi ve tanı doğruluk yüzdelerinin artırılması sağlanabilmektedir.

PE adaylarının tespitinde, akciğer damar ağacının doğru belirlenebilmesi için damar ağaç bölütlemesi ilk adım ve CAD sistem performansını etkileyen en önemli işlemdir [2]. Literatürde yapılmış çalışmalar incelendiğinde CAD sistemlerinin PE bölgelerini belirlemede belirli sınırlamalarla karşı karşıya kaldığı görülmektedir. Periferik emboliler gibi spesifik PE paternelerine göre dizayn edilmiş CAD sistemlerinin diğer tip emboli bölgelerinin belirlenmesinde başarısız sonuçlar ortaya çıkardığı görülmektedir [3]. Ayrıca sistem gürültüsü ele alınmadan sentetik veri setleri üzerinden tasarlanan CAD sistemlerinin gerçek CTA görüntülerinde oldukça kötü bir performans gösterdiği görülmektedir [4,5]. CTA görüntüleri üzerinde geliştirilen sistemlerde bir diğer sınır ise düşük emboli sayısına sahip veri setleri üzerinde geliştirilen CAD sistemlerdeki genelleme belirsizliğidir [6,7]. Bu çalışmalarda geliştirilen CAD sistemleri büyük veri setlerinde duyarlılık ve yanlış pozitif (YP)/veri seti oranları açısından kötü sonuçlar vermektedir.

PE adayları için elde edilen seçici özellikler incelendiğinde, bu konuda yapılan çalışmalarda PE adaylarının yoğunluk değerleri ve bu bölgelerin uzunluk hacim gibi yapısal özelliklerinin öne çıktığı görülmektedir [3-12].

Özellik çıkarma işlemi öncesinde gerçekleştirilen akciğer damar ağacı; eşikleme, BBE algoritması, damar hacimlerinin analizi gibi yöntemler ile gerçekleştirilen işlemler önceki çalışmalarımızda ele alınmıştır [13]. Bu çalışmada elde edilen PE adaylarından elde edilen belirleyici

Biyomedikal Sinyallerde sınıflandırma Uygulamaları

3. Gün 27 Eylül 2014 Cumartesi (09.45-10.45)

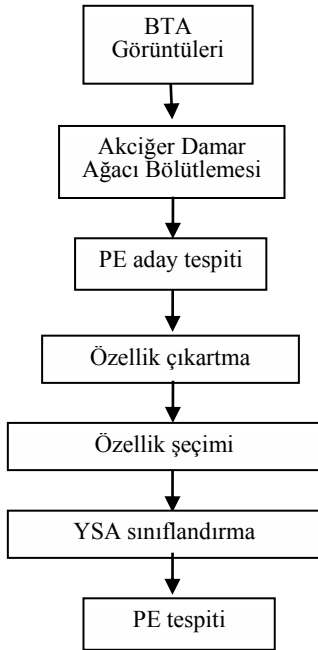
özellik matrisi ve küme içerisinde TBA ile çıkartılan indirgenmiş matrisler YSA ile karşılaştırılarak sonuçlar verilmiştir.

2. Metodoloji

Bu çalışmada BTA görüntülerinden PE'lerin tespit edilmesi için izlenen adımlar Şekil 1'deki gibidir. Çalışmanın anlaşılabilirliğini arttırmak adına ilk iki adım olan bölütleme ve PE aday belirleme işlemlerinden kısaca bahsedilecektir.

2.1 Akciğer Damar Ağacı Bölütlemesi

Akciğer damar bölümleri, başlangıç noktasından itibaren sırasıyla; Pulmoner trunk, arterler, lobar, segment ve subsegment damarlardır. Bölütleme işleminde ilk olarak, pulmoner trunk ve arterler birlikte bölütlenmiştir. İkinci olarak, lobar ve segmenter damarlar birlikte bölütlenmiş ve son olarak subsegmenter damarlar bölütlenmiştir. PE'ler ile birlikte bölütlenen alt dallar toplanarak akciğer damar ağacı elde edilmiştir [13,15].



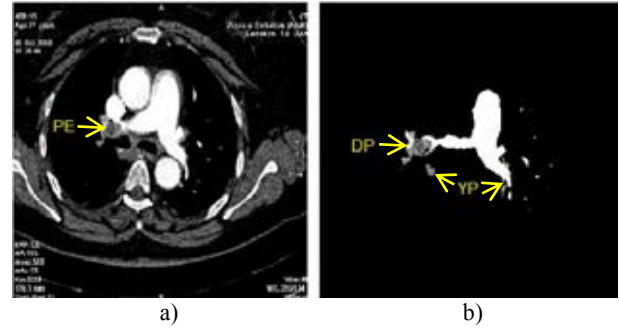
Şekil 1: Çalışmada takip edilen işlemler

Şekil 2a)' da bir hastaya ait BTA görüntülerinden 2 boyutlu bir kesit görülmektedir. Bu görüntüde radyologlar tarafından PE'nin yeri işaretlenmiştir. Şekil 2b)'de ise içerisinde PE olan bölütlenmiş 2 boyutlu akciğer damar ağacı görülmektedir [13].

2.2 PE Aday Tespiti

Elde edilen akciğer damar ağacı PE'leri içinde barındırmaktadır (Şekil 2b). Bu damarlardan PE adaylarının belirlenebilmesi için; önce içerisinde PE barındıran damar

ağacı bir maske olarak kullanılıp eşikleme işlemi uygulanmıştır. Ortaya çıkan görüntüye Bağlantılı Bileşen Etiketleme (BBE) algoritması uygulanarak, küçük boyutlu bileşenler görüntüden kaldırılarak, elde edilen son bileşenler PE adayı olarak tespit edilmişlerdir.



Şekil 2a)Orijinal BTA Görüntüsü b) Akciğer Damar Bölütlenmesi ve PE belirlenmesi

2.3 Özellik çıkartma

Sınıflandırıcı olarak kullanılacak akıllı sistemlerde, sistem başarısını arttırabilmek için sistem girişlerinin en uygun şekilde belirlenmesi gerekmektedir. PE adaylarında bulunan DP (Doğru Pozitif) ve YP (Yanlış Pozitif) bileşenleri adaptif olarak birbirinden ayırabilecek bir sistem kurabilmek için öncelikle bu iki bileşeni birbirinden ayırabilecek özelliklerinin belirlenmesi gerekmektedir. Bu işlem sayesinde DP ve YP bileşenlerin birbirlerinden farklılaşmaları sınıflandırma adımında başarı yüzdesini arttıracaktır. Örnek olarak BTA kesitleri boyunca uzanan hacimsel bileşenler ile tek bir kesitte bulunan 2-boyutlu bileşenleri birbirlerinden ayırmak için bileşenin var olduğu kesit sayısı bir özellik olarak düşünülebilir. Böylece hacimsel yapılar ile 2-boyutlu yapıları birbirinden ayırmak bu özelliğin kullanımı ile basitleşecektir.

Yapılan çalışmada elde edilen 14 ayrı özellik temel olarak üç başlık altında toplanabilir [15]. Çalışmada odaklanılan ilk adım PE adaylarının boyutsal uzunluk bilgileri olarak tanımlanabilir. Bu adımda PE adaylarının boyutsal uzunluklarının ortalama ve maksimum kullanılarak CTA görüntülerindeki gürültü bileşenlerinin YP olarak tanımlanabileceği görülmektedir. Bunun nedeni gürültü bileşenlerinin DP'lere göre daha kısa boyutlarda olmasıdır. Ayrıca PE adaylarının voksel sayıları bu başlık altında hesaplanmıştır.

Özellik seçiminde ikinci alt başlık PE adaylarının hacim ve alan bilgileridir. Alan ve hacim hesaplamalarında PE adaylarının boyut özelliklerinin en uzun olanı alan ve hacim elemanlarının kenar bilgisi olarak kullanılmıştır. Bu hesaplamalar yardımı ile olası en büyük yüzey ve hacimler elde edilerek PE adaylarının doldurabileceği dörtgen ve prizmatik uzaylar elde edilmiştir.

Üçüncü özellik alt kümesi ilk iki alt küme yardımı ile hesaplanan özelliklerin birbirleri ile olan oranları olarak tanımlanabilir. Yapılan çalışmada 33 hasta için elde edilen akciğer damar ağaçlarında bulunan PE adayları için çıkartılan özelliklerin listesi Tablo 1'deki gibidir.

Biyomedikal Sinyallerde sınıflandırma Uygulamaları

3. Gün 27 Eylül 2014 Cumartesi (09.45-10.45)

Tablo 1. PE adaylarından çıkartılan özellikler

	Çıkartılan özellikler
1	Maksimum kesit uzunluğu
2	Ortalama kesit uzunluğu
3	Maksimum dikey uzunluk
4	Ortalama dikey uzunluk
5	Maksimum yatay uzunluk
6	Ortalama yatay uzunluk
7	Voksel sayısı
8	Olası maksimum alan
9	Diagonal
10	Olası maksimum hacim
11	Voksel sayısı / Olası maksimum hacim
12	Voksel sayısı/ (Maks. dikey uz.+Maks.yatay uz)
13	Voksel sayısı/ (Maks. dikey uz.+Maks.yatay uz + Maks kesit uz.)
14	Voksel sayısı/ Diagonal

2.4 Özellik Seçimi

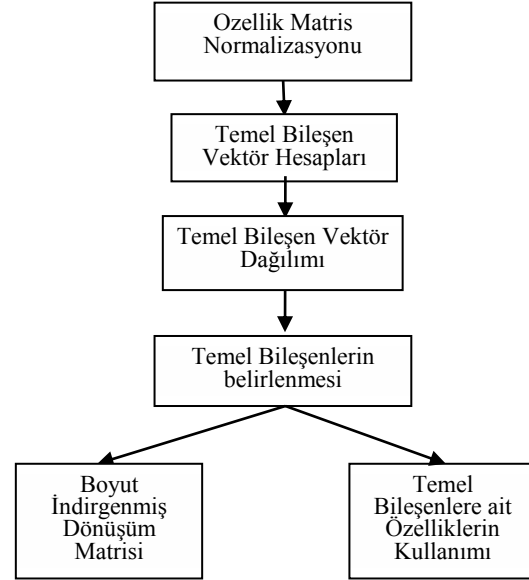
TBA, birbirleri ile ilişkiye (korelasyon) sahip özellik matrislerinde özellik sayısını indirmek için başvurulan tekniklerin en önemlilerindedir. Genel olarak özellikler arasındaki korelasyon yapısının yok edilip boyut indirgenmesi amacıyla kullanılmaktadır. Bu sayede aralarında korelasyon bulunan p sayıdaki özellik aralarında korelasyon olmayan ve sayıca orijinal özellik sayısından daha az sayıda ($p > k$) orijinal değişkenlerin (özelliklerin) doğrusal bileşenleri olan özelliklerle ifade edilmesidir.

TBA özellik matrisindeki değişim yapısı korelasyon ya da kovaryans matrisleri üzerinden incelenmektedir. Analizin hangi matris üzerinden yapılacağı değişkenlerde birim farklılığı olup olmamasına ve değişken varyanslarının yakın değer alıp almamasına bağlıdır. Temel bileşen hesabında birinci temel bileşen değişkeni varyansı en çok açıklayan en büyük özdeğere karşılık gelen özvektördür. İkinci temel bileşen ise birinci temel bileşenle korelasyonsuz değerlerin lineer kombinasyonunun varyansını maksimum yapan ikinci en büyük özvektördür.

Çalışmada kullanılan özellik çıkartım yöntemi Şekil 2'de gösterildiği gibidir. İlk aşamada uygulanan normalizasyon işlemi veri setini ölçekleme adına kullanılmış ve yapılan işlemler sırasıyla her bir gözlemi ortalama değerlerinden çıkartma ve standart sapmalarına bölme adımlarından oluşmaktadır. TBA için uygun hale getirilmiş özellik matris için ikinci adımda özdeğer ve özvektör değerleri bulunarak veri seti ile özvektörlerin çarpımından temel bileşen vektörleri hesaplanmıştır. Temel bileşen vektörlerinden oluşan matris için hesaplanan varyans (dağılım) değerleri temel bileşen vektörlerinin önem sıralamasının belirlenmesi adına oldukça önemlidir. Şekil 3'den de görüleceği gibi vektörlerin varyans değerlerinin kümülatif hesabı yapıldığında ilk 6 özelliğin özellik matrisinin %99.6'lık varyansını sakladığı ve 7-14 arasındaki özellikleri ilk altı özelliğin üzerine izdüşümünü

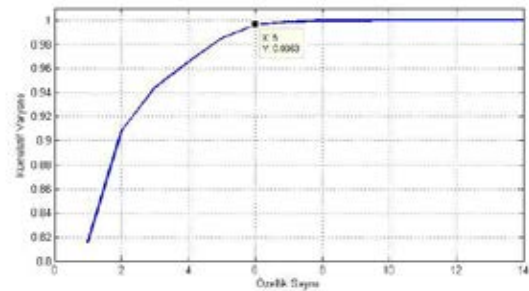
olarak %0.04'lük bir varyans kaybına neden olacağı görülmektedir.

Şekil 2: Özellik çıkartma işlemi



Özellik matrisi için çıkartılan temel bileşenler yardımı ile özellik indirgeme işlemi iki farklı yaklaşım ile yapılabilir. Bunlardan birincisi temel bileşenler yardımı ile geri dönüşüm yaparak boyut indirgemesi yapılmış yeni özellik matrisini kullanmak olarak tanımlanabilir. Bir diğeri ise temel bileşenlere sahip olan özellikleri ayırt edici özellik olarak belirlemektir.

Şekil 3: Temel bileşen varyans değerleri



2.5 YSA sınıflandırma

DP ve YP gibi farklı sınıfları birbirinden ayırmak için kullanılan sınıflandırıcı yapılarının en önemlilerinden biri YSA'dır. Çok katmanlı perseptron (MLP), girişine uygulanan verileri belirli çıkış değerlerine haritalayan, ileri yönlü, birden fazla katmana sahip eğitici bir yapay sinir ağıdır.

Çalışmada özellik matrisinin ve indirgenmiş iki farklı özellik seçim matrisinin sınıflandırılmasında kullanılan YSA yapısı MLP yapısıdır. 14 ve 7 nörondan oluşan iki adet gizli katmana sahiptir. Eğitim algoritması olarak Ölçekli Eşlenik Gardiyan (SCG) algoritması kullanılmıştır [14]. Ağ için momentum katsayısı 0.2, öğrenme katsayısı 0.02 olarak seçilmiştir. Eğitim ve test işlemleri için veri seti %70 ve %30 olarak rastgele bölünmüştür.



Biyomedikal Sinyallerde sınıflandırma Uygulamaları

3. Gün 27 Eylül 2014 Cumartesi (09.45-10.45)

2. Değerlendirme ve Sonuçlar

Bu çalışmanın sonucunda PE adayları için özellik çıkartım işleminin yapılması çıkartılan özellikler içerisinde özellik seçiminin gerçekleştirilmesi ve YSA ile sınıflandırma işlemleri sonunda elde edilen sonuçlar, tıbbi görüntü değerlendirmelerinde kullanılan performans ölçümleri ile analiz edilmiştir. Sonuç olarak, 33 hasta ve 450 PE durumu için, özellik çıkartımı ile elde edilen özellik matrisi, indirgenmiş özellik seçim matrisi ve temel bileşenlere sahip özelliklerin kullanıldığı özellik matrisi için sonuçlar Tablo 2'de verilmiştir.

Elde edilen sonuçlar eşliğinde en iyi sınıflandırma başarısı özellik çıkartımı ile elde edilen 14 özellikten oluşan matristir. TBA sonucunda elde edilen sınıflandırma başarısı temel özellik matrisine oldukça yakındır. Temel bileşen vektörlerini oluşturan ilk 6 özellik ile elde edilen sınıflandırma başarısı diğer iki sonuçtan daha düşüktür.

Çalışmada sonucunda TBA yardımı ile, korelasyona sahip veri setlerinde özellik indirgeme işleminin sınıflandırma başarısını yaklaşık olarak sabit tutarak sınıflandırma işlem hacmini düşürdüğü görülmüştür.

Tablo 2. YSA sınıflandırma sonuçları

Özellik Matrisi	Sınıflandırma Başarısı
Özellik çıkartım matrisi (14 özellik)	%84,99
TBA sonucu elde indirgenen özellik dönüşüm matrisi	%82,82
Temel bileşenlere sahip özellik matrisi (ilk 6 özellik)	%78,01

3. Kaynakça

- [1] R. Wittenberg, J. F. Peters, J. J. Sonnemans, M. Prokop, C. M. Schaefer-Prokop, Computer-assisted detection of pulmonary embolism: evaluation of pulmonary CT angiograms performed in an on-call setting, *Eur. Radiol.* 20 (4) (2010) 801–806.
- [2] ÖZKAN, H., Bilgisayarlı Tomografi Anjiyografi Görüntülerinde Pulmoner Embolilerin Bilgisayar Destekli Tespiti, Doktora tezi, Sakarya Üniversitesi, 2011
- [3] M. Das, A. Schneider, U. Schoepf, J. Cheema, S. Wood, P. Costello, et al., Computer-aided diagnosis of peripheral pulmonary emboli. *Radiological Society of North America* (2003) C02-232
- [4] S.R. Digumarthy, C.R. Kagay, A.C. Legasto, V.V. Muse, C. Wittram, J.O. Shepard, Computer-aided detection (CAD) of acute pulmonary emboli: Evaluation in patients without significant pulmonary disease., Chicago IL (2006) SSC04-08.
- [5] Z. Maizlin, P. Vos, M. Gody, P. Cooperberg, Computer-aided detection of pulmonary embolism on CT angiography: Initial experience, *J Thorac Imaging* 22 (4) (2007) 324-329.
- [6] Y. Masutani, H. MacMahon, K. Doi, Computerized detection of pulmonary embolism in spiral CT

- angiography based on volumetric image analysis, *IEEE Trans. Med. Imag* 21 (2002) 1517–1523.
- [7] E. Pichon, C.L. Novak, A.P. Kiraly, D.P. Naidich, A novel method for pulmonary emboli visualization from high-resolution CT images, In *Proc. SPIE Med. Imag* 5367 (2004) 161–170.
- [8] S. Buhmann, P. Herzog, J. Stoeckel, M. Salganicoff, M. Wolf, M. F. Reiser, et al., Clinical evaluation of a CAD prototype for the detection of pulmonary embolism, In *European Congress of Radiology* (2006) B-585.
- [9] C. Zhou, H.P. Chan, S. Patel, P.N. Cascade, B. Sahiner, L.M. Hadjiiski, et al., Preliminary investigation of computer-aided detection of pulmonary embolism in 3-D CT pulmonary angiography, *Acad. Radiol* 12 (6) (2005) 782–792.
- [10] A.P. Kiraly, C.L. Novak, D.P. Naidich, I. Vlahos, J.P. Ko, and G.T. Brusca-Augello, A comparison of 2D and 3D evaluation methods for pulmonary embolism detection in CT images, *Medical Imaging. Image perception, observer performance, and technology assessment. Proceedings of SPIE - Int'l. Society for Optical Engineering* (2006) 132-140. doi:10.1117/12.652430.
- [11] H. Bouma, J.J. Sonnemans, A. Vilanova, F. A. Gerritsen, Automatic Detection of Pulmonary Embolism in CTA Images, *IEEE Transactions on Medical Imaging* 28 (8) (2009) 1223-1230.
- [12] C. Zhou, L.M. Hadjiiski, B. Sahiner, H.P. Chan, S. Patel, P.N. Cascade, et al., Computerized detection of pulmonary embolism in 3-D CT images: Vessel tracking and segmentation techniques, In *Proc. SPIE Med. Imag* 5032 (2003) 1613–1620.
- [13] ÖZKAN H., OSMAN O., ŞAHİN S., “Bilgisayarlı Tomografi Anjiyografi Görüntülerinde Pulmoner Emboli Tespiti”, *TIPTEKNO 2012*, 12, 49-53, Antalya, Kasım 2012.
- [14] M. Moller, “A Scale Conjugate Gradient Algorithm for Fast Supervised Learning”, *Neural Networks* 6 (1993) 525-533.
- [15] ÖZKAN H., OSMAN O., ŞAHİN S., “BTA Görüntülerinde YSA ile Pulmoner Emboli Tespiti”, *TIPTEKNO 2013*, Antalya, 2013.

Teşekkür: Bu çalışma Dr. Siyami Ersek Göğüs Kalp ve Damar Cerrahisi Eğitim ve Araştırma Hastanesi tarafından desteklenmiştir. Hastane Başhekimi Sayın Prof.Dr. İbrahim Yekeler’e teşekkür ederiz.