

# Yeni Bir Temel/Bağımsız Bileşen Analizi(TBA/BBA) Tabanlı Öznitelik Seçme Yöntemi

## A New PCA/ICA Based Feature Selection Method

*Hakkı Murat Genç\* Zehra Çataltepe\*\* Thomas Pearson\*\*\**

\*Bilişim Teknolojileri Enstitüsü, Marmara Araştırma Merkezi, Kocaeli  
murat.genc@bte.mam.gov.tr

\*\*Bilgisayar Mühendisliği Bölümü, İstanbul Teknik Üniversitesi, İstanbul  
cataltepe@itu.edu.tr

\*\*\*Mühendislik Araştırmaları Birimi, USDA-ARS, Manhattan  
thomas.pearson@gmprc.ksu.edu

### Özetçe

Veri boyutunu azaltma yöntemleri sınıflandırma yapmak için harcanan zamanı ve bazı durumlarda sınıflandırma hatasını azaltmaya yardımcı olur. ([1], [2], [3], [4] ve [5]) Zaman kritik uygulamalarda öznitelik elde etme evresinde harcanan zamanı azaltmak için, öznitelik seçme yöntemleri, tüm giriş değerlerinin ölçülmesini gerektiren boyut indirgeme yöntemlerine tercih edilir. Geriye doğru öznitelik seçimi ve ileri doğru öznitelik seçimi gibi geleneksel yöntemlerin uygulanması hem çok maliyetlidir hem de en iyi olmayan öznitelik kümelerine neden olabilmektedir. Bu çalışmada, hangi özniteliklerin seçilmesi gerektiğine TBA veya BBA'nın hangi özniteliklere ne kadar değer atadıklarıyla karar verildiği yeni bir öznitelik seçme yöntemi tanıtılmaktadır. Yöntemin doğruluğu geriye doğru öznitelik seçimi, ileri doğru öznitelik seçimi ve aynı sayıda boyutun kullanıldığı TBA/BBA metodlarıyla karşılaştırılmıştır. Denemelerde mısır çekirdeklerinden elde edilmiş tayfsal ölçüm verisi kullanılmıştır.

### Abstract

Dimensionality reduction algorithms help reduce the classification time and sometimes the classification error of a classifier([1], [2], [3], [4] ve [5]). For time critical applications, in order to have reduction in the feature acquisition phase, feature selection methods are more preferable to dimensionality reduction methods, which require measurement of all inputs. Traditional feature selection methods, such as forward or backward feature selection, are costly to implement and may result in a suboptimal set of features. In this study, we introduce a new feature selection method that decides on which features to retain, based on how PCA (Principal Component Analysis) or ICA (Independent Component Analysis)[6] values those features. We compare the accuracy of our method to PCA and ICA using the same number of principal and independent components. We also do comparison to backward and forward feature selection with the same number of features selected. For our experiments, we use spectral measurement data taken from corn kernels infested and not infested by fungi.

### 1. Giriş

Öznitelik seçme yöntemleri örüntü tanıma uygulamalarında sıkça kullanılır. Bu yöntemler orijinal veri kümesinden belli sayıda öznitelik içeren bir alt küme seçerler. Öznitelik seçmenin önemli bir avantajı öznitelik elde etmek için harcanan zaman ve maliyetin azalmasıdır. Bunun yanında sınıflandırıcı eğitim ve test zamanları da kısalmır. Öznitelik seçimi gürlütlü, sapkın ve gereksiz özniteliklerin elenmesinde de yararlıdır.

Geriye ve ileri doğru öznitelik seçimi iyi bir öznitelik altkümüsi belirlenmesi için faydalı olabilir. Fakat bu yöntemler, her bir öznitelik eleme adımında bir sınıflandırıcının eğitilmesi ve test edilmesini gerektirdiğinden zaman bakımından oldukça maliyetlidir.

TBA(Temel Bileşen Analizi) ve BBA(Bağımsız Bileşen Analizi), boyut indirgemede sıkça kullanılırlar. Veri kümesi normal dağılıma yaklaştıkça TBA; normal dağılımdan uzaklaştıkça BBA en iyi performansına yaklaşır.

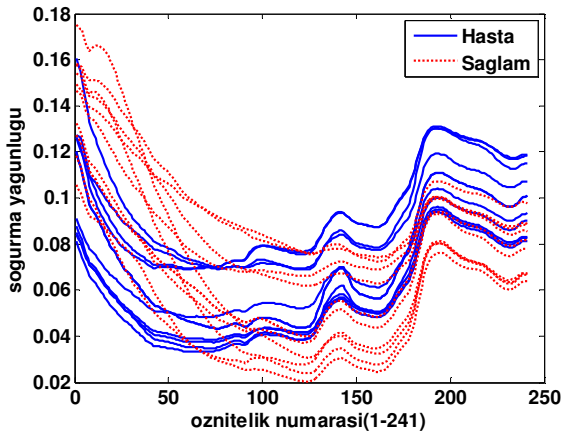
Bu çalışmada, TBA/BBA'nın özniteliklere atadığı değerlerin elenecek özniteliklerin seçimi için ölçüt olarak belirlendiği bir yöntem tanıtılmıştır. Önerilen yöntem tüm öznitelik kümesi ile başlayıp her bir adımda bir öznitelik elenmesi şeklinde çalıştığından geriye doğru öznitelik seçimi yöntemi ile benzerdir. Ama genellikle TBA/BBA uygulamak bir sınıflandırıcı eğitip test etmekten çok daha hızlı olduğundan tanıtılan yöntem geriye doğru öznitelik seçimi yönteminden çok daha hızlıdır. Önceki çalışmalardan Yu ve Liu[4],her öznitelik için 'ilgililik (relevance)' ve 'artıklık (redundancy)' değerini düzensizlik (entropy) tabanlı simetrik belirsizlikler kullanarak ölçmüştür. Yu ve Liu'nun önerdiği yöntem oldukça hızlı olmasına rağmen, en 'ilgisiz' olduğu sanılan her öznitelik elenmesinden dolayı bu özniteliklerin beraber kullanılması halinde getirebilecekleri fayda gözden kaçmaktadır. Simetrik belirsizlikler Hall'un[5] çalışmasında da öznitelikler arası ilişkilerin derecesini ölçmede kullanılmıştır.

Bildirinin geri kalan kısımları; veri kümesinin tanıtımı, genel kabul görmüş boyut indirme algoritmalarının kısaca özellikleri, ve yeni önerilen yöntemin tanıtımı şeklinde sıralanır. 5. bölümde her metodun sınıflandırma performansı doğruluk ve

zaman kıstasları açısından verilmiştir. 6.bölüm sonuçlar ve yöntemin gelecekteki açılımlarına ayrılmıştır.

## 2. Mısır Çekirdekleri Veri Kümesi

Kullanılan veri kümesi, çeşitli mantar türleri tarafından hastalık bulaştırılmış ya da bulaştırıl(a)mamış mısır çekirdeklerinden oluşur[7]. Pearson ve Wicklow'un tayfsal çekirdek yansıtma çarpanı, görünür renk yansıma imgeleri, x-ışını imgeleri, çok tayflı aktarlanlık imgeleri ve fiziksel özellikleri (kütle, uzunluk, kalınlık ve kesit alanı) de çekirdeğin hastalıklı olup olmadığını belirlemekte, hastalıklıysa buna neden olan mantar tipinin ayırt edilmesinde kullanılmıştır. Bu çalışmada ise sadece 550nm ile 1700nm arasında bulunan tayfsal çekirdek yansıtma çarpanları öznelikler olarak kullanılmış ve sadece hastalıklı olup olmama ayırımı yapılmıştır.(Problem 2 sınıfıdır.) Kullanılan veri kümesi, Pearson ve Wicklow'un çalışmalarında kullandıklarını içerir; fakat sınıflandırılması daha zor olan ortalama solgunluktaki yaklaşık 600 örneğin daha eklenmiş halidir. Sonuç olarak kullanılan veri kümesinde toplam 1648 mısır çekirdeği örneği ve her bir çekirdek için 550nm'den 1700nm'ye 5nm çözünürlükte elde edilmiş 241 farklı yansıtma çarpanı değeri bulunur. Veri kümesinin elde edilebilmesi için çekirdekler, haftalar öncesinden birçok farklı mantar türü ile aşılınmış mısır başaklarından hasat zamanı toplanmıştır. Daha sonra spektrometre ile bitkinin ışığı soğurma değerleri ölçülmüştür. Daha az sayıda dalga boyu için ölçüm almak kullanılan makinenin zamansal ve maddi maliyeti düşünülerek istenmektedir.



Şekil 1: Hastalıklı ve sağlam örneklerden rasgele seçilen bir kaç örneğin öznelik ölçüm uzayında dağılımı

Hastalıklı ve sağlam örneklerden bir grubu şekil 1'de sergilenmiştir. İlk bakışta 'kırmızı' ve 'mavi'lerin öznelik uzayının ilk ve son elemanlarında en ayırt edilebilir oldukları görülmektedir. Başka bir ifadeyle 1-15 arası (550nm-625nm) ve 200-241(1500nm-1700nm) arası öznelikler en iyi seçimlerdir. Veri kümesinin ortak değişim (covariance) matrisinin incelenmesinde de (yer darlığından dolayı verilmemiştir) en büyük köşegen girişlerin sırasıyla (200-241), (1-15) ve (175-185) numaralı özneliklere karşılık geldiği gözlenmiştir. Tablo 1'de en ayırt edici bölüm olan (200-241)den rasgele seçilen 5 ve 20 öznelik için doğruluk

değerleri, aynı sayıda temel bileşen için yapılan TBA sonuçları ile karşılaştırılmıştır.

Tablo 1: Fisher'in doğrusal sınıflandırıcısı için ortalama geçişleme hatalarının karşılaştırılması

Yöntem ⇒ ↓ #öznelik/bileşen	TBA	(200-241) numaralı özneliklerden rasgele seçim	(200-241) numaralı özneliklerden rasgele seçim(önişlemlenmiş)
5	0.1181	0.1205	0.1145
20	0.0860	0.0804	0.0764

Birçok uygulamada, öznelik seçme/boyut indirgeme işleminden önce veriyi önişlemlerden geçirmek uygundur. Elimizdeki mısır veri kümesi tayfsal bilgi içerdiğinden ve öznelikler komşularına bağımlı olduğundan düzgeleme (normalization) işlemi gerekmemekte fakat yumuşatma işlemi önemli olabilmektedir. Yumuşatmanın, başka bir ifade ile her değer için komşu öznelik değerlerinin de ağırlıklı ortalamasını alma işleminin bir diğer önemli yararı ise sapkın(istisna) öznelik değerlerinin elenmesidir. Mısır çekirdeği veri kümesinde ortalamadan önemli ölçüde sapan değerlere fazla rastlanmadığından bu kriter ikinci önemdedir. Yumuşatma işleminin olumlu etkileri 4 ve 5 nolu tabloların 2 ve 3 nolu tablolarla karşılaştırmalı olarak incelenmesi ile gözlemlenebilir.

## 3. Öznelik Seçme/Boyut İndirgeme Algoritmaları

Geriye doğru öznelik seçimi, ileri doğru öznelik seçimi TBA ve BBA([6] ve [8]) sıklıkla kullanılan öznelik seçme ve boyut indirgeme algoritmalarıdır[9].

Geriye doğru öznelik seçimi ve ileri doğru öznelik seçimi, her özneliğin, seçilecek öznelik alt kümesine eklenmesine ya da bu kümeden çıkarılmasına her adımda bir sınıflandırıcının eğitilip test edilmesi ile karar verilen 'sarıcı' yöntemlerdir. Çok geniş veri kümeleri için bu yöntemlerin uygulanması olanaksızlaşır. İleri doğru öznelik seçimi algoritması yalnız başına değersiz ama birlikte oldukça değerli olabilecek öznelikleri kaçırabilir.

TBA ve BBA  $d$  boyutlu giriş uzayını  $d' < d$  olacak şekilde farklı bir uzaya aktarır. BBA'da amaç normal dağılımlı olmayan verinin istatistiksel olarak bağımsız ya da olabildiğince bağımsız olmasını sağlayacak şekilde bir doğrusal dönüşümünü elde etmektir. TBA, verinin normal dağılımlı olmasından yararlanırken BBA sınıflandırıcı performansını normal dağılımdan uzaklaştığı ölçüde artırabilir. Bu yüzden bu iki yöntemle denemeler yapıldığında verinin istatistiksel dağılımına ilişkin öngörülerin her iki uç noktaya denemeler olur. Mısır çekirdeği veri kümesi bu açıdan değerlendirildiğinde iki yöntem için de benzeri sonuçlar vermektedir. (Bkz. Tablo 2 ,3, 4, 5)

## 4. TBA/BBA Tabanlı Öznelik Seçme Algoritması

Hem TBA hem BBA tabanlı öznelik seçme algoritmaları, temel ya da bağımsız bileşenlere en az katkı yapan özneliklerin en az ayırt edici özelliğe sahip olduğu fikrine

dayanır. Her aşamada en 'kötü' öznitelik elenir ve kalanlar üzerinden tekrar TBA/BBA yapılır. Özniteliklerin bileşenlere katkısı, dönüşüm matrisi  $W$ 'nin ilgili satır ya da sütunundaki elemanların mutlak değerlerin toplamı olarak yakınsanmıştır.

$d$  giriş veri boyutu,  $d'$  TBA/BBA ile elde edilen dönüştürülmüş veri kümesinin boyutu ve  $d^*$  istenilen öznitelik boyutu olarak tanımlansın. Başlangıçta  $d$  değeri orijinal veri kümesi boyutumuzdur. Takip eden algoritma adımlarını sıralayacak olursak,

- Temel/bağımsız bileşenleri hesapla.  $W$  dönüşüm matrisini elde et.  $A_{N \times d}$  orijinal veri kümesi,  $Z_{d \times d}$  elde edilen dönüşüm ve  $W_{d \times d}$  dönüşüm matrisi olmak üzere  $Z = AW$  yazılabilir.
- $W$  matrisinin tüm satır elemanlarının mutlak değerini ayrı ayrı topla.  $W$  matrisinin her bir satırı farklı bir öznitelikle çarpılacağından öznitelikler dönüşüm için birer katsayı kümesi olarak düşünülebilir. Eğer belirtilen satır toplamı 'küçük'se, ilgili özniteliğin temel/bağımsız bileşenlere katkısı 'küçük'tür.
- En küçük satırlar toplamını bul ve ilgili özniteliği ele. ( $d \rightarrow d - 1$ )
- Eğer  $d > d^*$  koşulu geçerliyse(hala istenilenden fazla öznitelik varsa), geriye kalan öznitelikleri kullanarak 1. adıma dön.

TBA'da özdeğerlerin büyüklükleri de dikkate alınarak istenilen sayıda temel bileşene ( $d'$ ) ulaşılmaya çalışılabilir. Buradaki tek kısıt istenen öznitelik sayısının bileşen sayısından büyük ya da eşit olmasıdır. ( $d' \leq d^*$ ) Örneğin, 241 öznitelikten 5 'en iyi' özniteliği seçmek için en çok 5 temel bileşen için analiz yapılmalıdır. BBA'nde ise her adımda bağımsız bileşen sayısı algoritma tarafından en iyilenmektedir ve benzeri bir kısıt algoritmanın doğası gereği sağlanmaktadır.

## 5. Performans Sonuçları

### 5.1. Kullanılan sınıflandırıcılar

Önerilen algoritmanın sonuçlarını karşılaştırmak için Fisher'in doğrusal sınıflandırıcısı ve lojistik doğrusal sınıflandırıcı kullanılmıştır. ([9],[10] ve [11])

Fisher'in doğrusal sınıflandırıcısı, en küçük kareler bazında hatayı en azaltacak doğrusal ayırtaç fonksiyonunu bulmaya çalışır. Lojistik doğrusal sınıflandırıcı ise olabilirlik ölçütünü lojistik (sigmoid) fonksiyonunu kullanarak en çoklayarak veri kümesi için bir doğrusal sınıflandırıcı hesap etmeye çalışır [10, 12].

### 5.2. Deneysel Sonuçlar

Sonuçlar, orijinal veri kümesi için tablo 2 ve 3'te, önişlemeden geçirilmiş veri kümesi için tablo 4 ve 5'te raporlanmıştır. Tüm sonuçlar 10'lu çapraz geçirme yöntemi ile elde edilen verinin ortalaması olarak verilmiştir.

Önişlemeden geçirilmiş veri kümesi, orijinalin yarı maksimum tam genişlik(FWHM) bant geçiren süzgeçle evrişiminden elde edilmiştir. Veri yüksek ilintili olduğundan ve bu nedenden düzgeleme veri kaybına neden olabileceğinden, tüm durumlarda veri 0 ortalamalı ve 1 değışintili olacak şekilde düzgelememiştir.

Önerilen yöntem özünde bir geriye doğru öznitelik seçimidir. Sonuçlar incelendiğinden özellikle düşük sayıda öznitelik seçiminde yöntemlerin, klasik geriye/ileriye doğru öznitelik seçimi yöntemlerinden iyi doğruluk değerlerine sahip olabildiği görünmektedir. Fakat bu klasik yöntemleri çok geniş veri kümelerinde kullanılamaz.

Klasik TBA ve BBA yöntemleri ile aynı sayıda bileşen kullanıldığında önerilen yöntemler kadar iyi sonuç alınmadığı görülmektedir. Ayrıca TBA ve BBA birer dönüşüm yöntemi olduğundan tüm öznitelik değerlerini kullanır. Bu durum ise gerçek zamanlı mısır çekirdeği sınıflandırma probleminde öznitelik elde edilmesi aşamasında ciddi zaman ve para kaybına neden olduğundan tercih edilmez.

Tablo 2: Fisher'in doğrusal sınıflandırıcısı için ortalama geçirme hataları(Orijinal veri kümesi)

Yöntem $\Rightarrow$ #öznitelik/bileşen $\downarrow$	Geriye Doğru Öznitelik Seçimi	İleri Doğru Öznitelik Seçimi	TBA tabanlı Öznitelik seçimi	BBA tabanlı Öznitelik seçimi	TBA	BBA
1	0.2323	0.2049	0.2230	0.2121	0.2751	0.3836
3	0.1445	0.0945	0.1333	0.1261	0.1175	0.3896
5	0.0994	0.0689	0.1181	0.1012	0.1181	0.4169
10	0.0360	0.0384	0.0896	0.0762	0.1139	0.3103
20	0.0079	0.0213	0.0490	0.0570	0.0860	0.1430
30	0.0049	0.0110	0.0357	0.0285	0.0309	0.1121

Tablo 3: Lojistik doğrusal sınıflandırıcı için ortalama geçirme hataları(Orijinal veri kümesi)

Yöntem $\Rightarrow$ #öznitelik/bileşen $\downarrow$	Geriye Doğru Öznitelik Seçimi	İleri Doğru Öznitelik Seçimi	TBA tabanlı Öznitelik seçimi	BBA tabanlı Öznitelik seçimi	TBA	BBA
1	0.2043	0.2049	0.2218	0.2152	0.2690	0.3890
3	0.1122	0.0902	0.1654	0.1339	0.1430	0.4084
5	0.0409	0.0573	0.1224	0.0976	0.1363	0.4157
10	0.0238	0.0238	0.0757	0.0745	0.0963	0.3042
20	0.0134	0.0128	0.0369	0.0558	0.0624	0.1333
30	0.0165	0.0085	0.0175	0.0285	0.0236	0.0866

Tablo 4: Fisher'in doğrusal sınıflandırıcısı için ortalama geçirme hataları(Önişlemeden geçirilmiş veri kümesi)

Yöntem $\Rightarrow$ #öznitelik/bileşen $\downarrow$	Geriye Doğru Öznitelik Seçimi	İleri Doğru Öznitelik Seçimi	TBA tabanlı Öznitelik seçimi	BBA tabanlı Öznitelik seçimi	TBA	BBA
1	0.2305	0.2049	0.2248	0.2121	0.2757	0.3824
3	0.1530	0.0939	0.1309	0.1055	0.1169	0.3848
5	0.0884	0.0695	0.1145	0.0875	0.1181	0.3975
10	0.0421	0.0396	0.0896	0.0679	0.1157	0.3654
20	0.0110	0.0189	0.0424	0.0364	0.0921	0.1800
30	0.0055	0.0091	0.0212	0.0273	0.0200	0.0800

Tablo 5: Lojistik doğrusal sınıflandırıcı için ortalama geçirme hataları(Önişlemeden geçirilmiş veri kümesi)

Yöntem $\Rightarrow$ #öznitelik/bileşen $\downarrow$	Geriye Doğru Öznitelik Seçimi	İleri Doğru Öznitelik Seçimi	TBA tabanlı Öznitelik seçimi	BBA tabanlı Öznitelik seçimi	TBA	BBA
1	0.2049	0.2049	0.2218	0.2152	0.2690	0.3860
3	0.1201	0.0890	0.1563	0.1048	0.1424	0.3884
5	0.0494	0.0598	0.1109	0.0936	0.1363	0.4060
10	0.0244	0.0250	0.0842	0.0739	0.0957	0.3666
20	0.0128	0.0116	0.0254	0.0455	0.0624	0.1678
30	0.0220	0.0116	0.0090	0.0267	0.0727	0.0703

Tablo 6 ve 7’de, önerilen yöntemlerin hızı, klasik geriye doğru öznelik seçimi yöntemi ile karşılaştırılmıştır. (Sonuçlar her bir durum için 3.2 P4 dual core işlemcili, belleği 2Gb olan adanmış bir bilgisayarda elde edilmiştir.). Önerilen yöntemlerin hızlarının, kullanılan sınıflandırma metodu ile değişmediği bu tablolardan gözlemlenebilir. Bu özellik, sınıflandırmada yapay sinir ağları, karar destek makineleri gibi daha fazla zaman alan yöntemler kullanılması durumunda önerilen öznelik seçme algoritmasının önemini daha da artırır.

Tablo 6: Fisher’ın doğrusal sınıflandırıcısının öznelik seçimi için harcadığı zaman(saniye cinsinden)

Yöntem ⇒ #öznelik/bileşen ↓	Geriye doğru öznelik seçimi	TBA tabanlı öznelik seçimi	BBA tabanlı öznelik seçimi
3	7886	182	2834
5	7885	182	2834
10	7883	182	2833
20	7876	182	2828
30	7864	181	2820

Tablo 7: Lojistik doğrusal sınıflandırıcısının öznelik seçimi için harcadığı zaman(saniye cinsinden)

Yöntem ⇒ #öznelik/bileşen ↓	Geriye doğru öznelik seçimi	TBA tabanlı öznelik seçimi	BBA tabanlı öznelik seçimi
3	52328	182	2835
5	52327	182	2835
10	52324	182	2834
20	52314	182	2829
30	52290	182	2821

## 6. Sonuçlar ve Gelecek Çalışmalar

TBA/BBA tabanlı hızlı ve doğru bir geriye doğru eleme algoritması önerilmiştir. Yöntem benzeri klasik yöntemlerle karşılaştırılmış ve sonuçlar raporlanmıştır.

Her iki yöntemin de seçtiği öznelikler, veri kümesinin incelendiği 2. bölümde öngörülen en çok ayırt edici kesimlerdenidir. Ayrıca sonuçlar tablo 1’de raporlanan sonuçlardan daha iyidir.

Algoritmanın en önemli kısıdı, negatif değere sahip olmayan veri kümelerinde yüksek doğrulukla çalışabilmesidir. Yöntem bu özelliğe sahip farklı bir veri kümesinde de denenmiş ve benzeri sonuçlar elde edilmiştir. (OpdtDigit veri kümesi [13]) Sonuçlar yer yetersizliğinden verilmemiştir. Sadece tayfsal (spectral) bilgi taşıyan veri kümelerinde yumuşatma işlemi algoritma performansını artırır.

Önerilen yöntem her bir adımda birden fazla özneliğin elenmesi yoluyla hızlandırılabilir. Sınıflandırıcı doğruluğundan ödün vermek kaydıyla, bileşenler en ‘büyük’ ‘katkı’ yapan istenen sayıda öznelik ilk adımdan seçilebilir.

Yöntemin çok büyük veri kümelerinde klasik yöntemlere göre kazandırdığı zamanın irdelenmesi de gelecek çalışmalarımız arasındadır.

## 7. Teşekkür

Yazarlar, deneysel sonuçların elde edilmesinde kullanılan PRTools4[14] ve FastICA[8] araç takımı yazarlarına ve doktora çalışmalarına sağladığı maddi kaynaktan ötürü TÜBİTAK’a teşekkürlerini sunar.

## 8. Kaynakça

- [1] Y. Yang and J. Pedersen, “A comparative study on feature selection in text categorization,” in *International Conference on Machine Learning*, 412420, 1997.
- [2] J. Bins and B. A. Draper, “Feature selection from huge feature sets,” in *8<sup>th</sup> International Conference on Computer Vision (ICCV 01)*, 159165, 2001.
- [3] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” *Journal of Machine Learning Research*, 3, 11571182, 2003.
- [4] L. Yu and H. Liu, “Efficient feature selection via analysis of relevance and redundancy,” *Journal of Machine Learning Research*, 5, 12051224, 2004.
- [5] M. A. Hall, “Correlation-based feature selection for discrete and numeric class machine learning,” in *the Seventeenth International Conference on Machine Learning*, 159-165, 2000.
- [6] A. Hyvarinen and E. Oja, “Independent component analysis: Algorithms and applications,” *Neural Networks*, 13, 411430, 1999.
- [7] T. C. Pearson and D. T. Wicklow, “Properties of corn kernels infected by fungi,” *Transactions of the ASAE*, 49, 12351245, 2006.
- [8] A. Hyvärinen. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks* 10(3):626-634, 1999.
- [9] R. Duda, P. Hart and D. Stork, *Pattern classification*, 2nd ed. New York: John Wiley and Sons, 2001.
- [10] A. Webb, *Statistical pattern recognition*. New York: Wiley, 2002.
- [11] S. Raudys and R. Duin, “On expected classification error of the fisher linear classifier with pseudoinverse covariance matrix,” *Pattern Recognition Letters*, 19, 385392, 1998.
- [12] E. Alpaydin, *Introduction to machine learning*, The MIT Press, Cambridge, Massachusetts, 2004.
- [13] E. Alpaydin, C. Kaynak, *Optical Recognition of Handwritten Digits*, Department of Computer Engineering, Boğaziçi University, 80815 Istanbul Turkey, July 1998
- [14] R.P.W. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. de Ridder, D.M.J. Tax, *PRTools4, A Matlab Toolbox for Pattern Recognition*, Delft University of Technology, 2004.