

Denetimli Örüntü Tanıma ve Gıda Analizlerinde Uygulamaları

Bahar Demircan , Yeşim Elmacı  

Ege Üniversitesi, Mühendislik Fakültesi, Gıda Mühendisliği Bölümü, 35100 Bornova, İzmir

Geliş Tarihi (Received): 13.07.2018, Kabul Tarihi (Accepted): 27.09.2019

✉ Yazışmalardan Sorumlu Yazar (Corresponding author): yesim.elmaci@ege.edu.tr (Y. Elmacı)

☎ 0 232 311 13 16 📠 0 232 342 75 92

ÖZ

Denetimli örüntü tanıma, sınıflandırma için örnek kategorisi üyeliği hakkında bir ön bilginin kullanıldığı teknikleri ifade etmektedir. Sınıflandırma modeli, kategorileri olan örneklerin bir eğitim seti üzerinde geliştirilmektedir. Kimya, biyoloji, ilaç ve gıda bilimi içinde denetimli örüntü tanıma uygulaması giderek daha önemli hale gelmektedir. Denetimli örüntü tanıma yöntemleri çok çeşitlidir ve asıl önemli nokta en uygun yöntemi seçmektir. Gıda analizlerinde gıda kalite değerlendirmesi, veri yorumlama gibi çeşitli amaçlarla farklı verilere uygulamaları bulunmaktadır. Denetimli örüntü tanıma teknikleriyle incelenen gıdalara örnek olarak şarap, yağ, bal, süt ürünleri, et, meyveler, içecekler, tahıllar ve balık verilebilir. Bu teknikler kullanılarak gıdalarda doku analizi, aroma analizi, gıda doğrulaması, gıda kalitesinin değerlendirilmesi, çoklu element analizi, coğrafi ve botanik kökene göre sınıflandırma gerçekleştirilebilmektedir. Bu derlemede, denetimli örüntü tanıma tanımlanmış, uygulama teknikleri özetlenmiş ve gıda analizlerinde kullanılan örüntü tanıma teknikleri konusunda yapılan çalışmalar ile örneklendirilerek bilgi verilmiştir.

Anahtar Kelimeler: Denetimli örüntü tanıma, Gıda analizi, Çok değişkenli veri analizi, Görüntü işleme, Sınıflandırma

Supervised Pattern Recognition and its Applications in Food Analyses

ABSTRACT

Supervised pattern recognition is a technique for classification that the prior knowledge is used regarding member of sample category. Classification model is improved by using samples separated into category as training set. Supervised pattern recognition is getting more important for chemistry, biology, pharmacology and food science. There are many supervised pattern recognition methods. The main part is to select the most appropriate method. There are implementations to different inputs for various purposes such as food quality assessment and data interpretation. Wine, oil, honey, dairy products, meat, fruits, beverages, cereals and fish could be given as examples analyzed by supervised pattern recognition techniques. Also by using this techniques, texture and aroma analyses, food verification, food quality assessment, multiple element analysis, classification based on geographical and botanical origins can be performed. In this review, supervised pattern recognition is defined, its application techniques are summarized, and information is provided by exemplifying studies on pattern recognition techniques used in food analysis.

Keywords: Supervised pattern recognition, Food analysis, Multivariable data analysis, Image processing, Classification

GİRİŞ

Günümüzde, modern analitik araçlar, nispeten kısa sürede analiz edilebilecek çok sayıda örnek (nesne) için çok miktarda bilgi (değişken veya özellik) üretmeye izin vermektedir. Bu durum verilerin maksimum yararlı bilgisini verimli bir şekilde çıkarmak için matematiksel ve istatistiksel yöntemlerin kullanılmasını gerektiren çok değişkenli veri matrislerinin kullanılabilirliğine yol açmaktadır. Denetimli örüntü tanıma teknikleri, bilinen sınıflardan birinde bilinmeyen örneklerin ölçüm modeline göre sınıflandırılması için örneklerin sınıf üyeliği hakkındaki bilgileri belirli bir gruba dahil etmede (sınıf veya kategori) kullanılmaktadır [1, 2]. Denetimli örüntü tanıma prosedürleri, algoritmanın uygulandığı, aşağıdaki adımlardan oluşan ortak bir strateji kullanarak yürütülmektedir:

- I. Değişkenlerin ölçüldüğü bilinen sınıf üyeliği nesnelere ilişkin bir eğitim, kalibrasyon ve test seti seçimi,
- II. Değişken seçimi,
- III. Eğitim setini kullanarak model oluşturulması,
- IV. Elde edilen sınıflamanın güvenilirliğini değerlendirmek için, bu yöntemin geçerliliğinin onaylanması [2, 3].

Gıda bilimlerinde birkaç çeşit örüntü tanıma yöntemi uygulanmış olup bunlar, sınıflandırmaya ulaşma yolunda esas olarak farklılık göstermektedir [2]. İki tip yöntem genellikle ilk yaklaşımda ayrılır. Bu ayrım doğrusal tanımlayıcı analiz (LDA), k-en yakın komşular (kNN), sınıflandırma ve regresyon ağaçları (CART), kısmi en küçük kareler tanımlayıcı analiz (PLS-DA) gibi sınıflar arasında sınıflandırma üzerine odaklananlar ile yapay sinir ağları (YSA), sınıf analogunun esnek bağımsız modellemesi (SIMCA), eşit olmayan dağıtık sınıfların esnek bağımsız modellemesi (UNEQ) gibi modelleme sınıflarına yönelik olanlar şeklinde [4]. Sınıflandırma teknikleri, sınıflandırma ile ilgili tüm kategorilere dayalı modeller oluşturmak için kullanılırken, ayrık sınıf modelleme yöntemleri, her bir kategori için ayrı bir model oluşturmaktadır. Ayırıcı yöntemlerin dezavantajlarından biri, örneklerin, bunlardan herhangi birine ait olmaları bile, her zaman verilen kategorilerden biri olarak sınıflandırılmasıdır. Sınıf modelleme yöntemleri, modelin bir parçası olarak modele uygun olan nesnelere göz önünde bulundurmada ve olmayan üyeler olarak sınıflandırmaktadır. Ancak, sınıflandırma yetenekleri ile ilgili olarak, yakın zamandaki ampirik inceleme, SIMCA'nın daha önce düşünüldüğü kadar güçlü olmadığını göstermektedir; sınıflandırma ve etki matrisi analizi (CAIMAN), CART ve tanımlayıcı analizin varyantlarının daha iyi olduğu ifade edilmektedir [4, 5].

Denetimli örüntü tanıma teknikleri parametrik/parametrik olmayan [2], belirleyici/olasılıksal veya doğrusal/doğrusal olmayan yöntemler olarak gruplandırılabilir. LDA, PLS-DA, SIMCA ve UNEQ gibi parametrik teknikler, karar fonksiyonunun türetilmesinde nesnelere dağılımının istatistiksel parametrelerini kullanmaktadır (genellikle çok değişkenli normal dağılım varsayılır). KNN ve CART gibi

parametrik olmayan yöntemlere ait istatistikler, doğru sınıflandırma olasılıklarının tahminini zorlaştıran dağıtım varsayımına dayanmamaktadır. Doğrusal/doğrusal olmayan sınıflandırma, sınıflar arasında ayrım yapmak için kullanılan, sırasıyla doğrusal/doğrusal olmayan tanımlayıcı fonksiyonlarının doğasına dayanır [6].

Bu derlemede günümüzde otomotiv, astronomi, tıp ve güvenlik gibi sektörlerde yaygın olarak kullanılan denetimli örüntü tanıma tekniklerinin gıda analizlerinde kullanımı hakkındaki çalışmalar incelenmiş ve söz konusu tekniklerin gıda analizlerinde doğru bir şekilde kullanımının pratik ve faydalı sonuçlar elde etmeye olanak sağlayabileceği düşünülmüştür. Derlemede kapsamlı olarak denetimli örüntü tanıma uygulamasının amacı, teknikleri, işlem basamakları ve sonuçların yorumlanması gıda analizlerinde kullanım örnekleriyle birlikte incelenmiştir.

DENETİMLİ ÖRÜNTÜ TANIMA

Nesnelerin ve olayların tespiti ve sınıflandırılması olarak tanımlanan örüntü tanıma, özellikle nesnelere makinalar ile kategorilere ayrılmasını ifade eder. Ölçüm ve nitelikler, nesnelere sınıflandırılmasında özellik olarak adlandırılırken, tip ve kategoriler ise nesnelere gruplandırılmasında sınıf olarak adlandırılmaktadır. Sınıflandırmadaki bireysel parçalar nesne olarak, durumlar da örnek veya örüntü olarak adlandırılmaktadır. Denetimli (eğitilmiş) öğrenme; eğitime uygulamalarında bulunan giriş vektörlerinin ve bunlara uyan hedef vektörlerinin verildiği örnekler olarak tanımlanırken, denetimsiz (eğitici) öğrenme ise örüntü tanıma problemlerinde eğitim verisinin herhangi bir hedef vektörü olmadan bir x giriş vektörü dizisinden oluşması olarak tanımlanmaktadır. Endüstride kalite kontrol, işlem kontrolü ve nesne tanımda kullanılan örüntü tanıma uygulamaları otomotiv, astronomi, biyoloji, tıp, finans ve güvenlik gibi alanlarda kullanılmaktadır. Bu uygulamalar son dönemlerde gıda analizlerinde de kullanılmaya başlanmıştır [7]. Günümüzde en yaygın kullanılan örüntü tanıma uygulamasının örnekleri:

- Tıbbi görüntülerin, fotoğrafların otomatik olarak analiz edilmesi,
- İnsan konuşmasının bilgisayarla tanınması,
- Petrol ve mineral araştırmaları ve deprem tespiti için sismik sinyallerin sınıflandırılması,
- Parmak izinden, el şeklinden, retinadan, ses karakteristiğinden, el yazısından kimlik tespiti,
- Çizilmiş desenlerin ve basılmış karakterlerin otomatik olarak tespiti ve el yazısı tanıma,
- Hava durumu, kar ve su rezervleri, mineralleri tespit etmek için kullanılan uydu görüntülerinin analizi,
- Montaj bandındaki parçaların otomatik denetimi,
- Kalp hastalarının tıbbi kategorilerde sınıflandırılması, dalgaların tespiti ve analizi olarak sıralanabilmektedir [7]

Bu örnekler uygulama alanı, uygulama, giriş örüntüsü ve örüntü sınıfları şeklinde detaylandırılabilir. Bu şekilde günümüzde yapılan çalışmalara:

- Optik karakter tanıma uygulamalarının doküman analizinde karakter ve kelimelerin sınıflandırılmasında giriş örüntüsü olarak doküman görüntüsünün kullanılması,
- Tıpta teşhis amacıyla kanserli ve sağlıklı hücrelerin sınıflandırılmasında giriş örüntüsü olarak mikroskobik görüntülerin kullanılması,
- Otomatik hedef tanıma amacıyla askeriyede hedef tiplerinin sınıflandırılmasında giriş örüntüsü olarak optik ya da infrared fotoğrafların kullanılması,
- Giriş örüntüsü olarak yürüyen bant üzerinden alınan meyve görüntülerinin, meyve sınıflandırma uygulamalarıyla endüstriyel otomasyonda kalite derecelerine göre sınıflandırılması
- DNA dizilerinin giriş örüntüsü olarak kullanıldığı biyoenfeksiyon uygulamalarında dizi analizleri amacıyla bilinen gen tiplerine göre sınıflandırma yapılması örnek olarak verilebilir [7].

Örüntü Tanıma Süreci

Örüntü tanıma sürecinin ilk basamağında algılama işlemini gerçekleştirmek için fiziki çevreden veriler toplanmakta, ardından bir ön işleme basamağı ile özellik çıkarım işlemleri gerçekleştirilmektedir. Sınıflandırmak üzere özellikler belirlenmekte ve bu süreçle birlikte eğitim verilerinin özellikleri de ön işleme ve özellik çıkarma basamaklarından geçirilerek belirlenmektedir. Alt yapı oluşturduktan sonra seçilen özellikler, model öğrenmede elde edilen model ile diğer özelliklerin sınıflandırılmasında kullanılmaktadır. Karar verme basamağında ise sonuçlar üzerinde gerekli düzenlemeler yapılarak işlem sonlandırılmaktadır [8].

Sembolik (örnek: renk) ve nümerik (örnek: uzunluk) olarak belirtilen "öznitelik", herhangi bir ayırt edici durum, nitelik ya da karakteristik özellik olarak tanımlanabilir. "Örüntü" ise bireye ait karakteristik özellik ya da özniteliklerin bileşimi olarak tanımlanmaktadır. Öznitelik vektörü X , gözlemler ya da özniteliklerden oluşurken, W ise gözlemlerin arkasındaki kavram veya etiket olarak ifade edilmektedir. Örüntü " $\{X, W\}$ " şeklinde gösterilmekte, sınıflandırma ise öznitelik uzayını sınıf etiketli karar bölgelerine ayırmak olarak tanımlanmaktadır. Sonuç olarak, öznitelik vektörü olarak ifade edilen X hangi karar bölgesine ait ise örnek o sınıfa atanmaktadır [8].

Görüntü İşleme Teknikleri

Son yıllarda kullanımı oldukça artan görüntü işleme ve bilgisayar kullanımı ile görüntü tanıma uygulamalarının özellikle araç içi otomasyon, güvenlik sistemleri, gezgin robot uygulamaları, askeri alanlarda dost ve düşman kuvvetlerinin gözetilmesi, tarım uygulamaları, biyomedikal ve tıp alanlarında, coğrafi bilgi sistemlerinde, tasarım ve imalat uygulamaları gibi geniş kullanım alanları bulunmaktadır [9, 10].

Görüntü işleme teknikleri uygulamalarında önce kameradan görüntüler alınmakta ve bu görüntüler üzerinde görüntü ön işleme basamakları gerçekleştirilerek ilgili nesnelere ait özellik çıkarımı yapılmaktadır. Bu aşamada ortamdaki nesnelere doğru

tanınması önem taşımaktadır. Nesne tespit veya tanınmasında farklı yöntemlerin kullanımı önerilmektedir. Bu amaçla hızlı ve etkili sonuç almak için nesnelere basit özellikleri kullanılarak yapılan çalışmalarda [11], karmaşık arka plan çıkarımı ile tanıma [12], şekil tanıma, renk tanıma, kenar ve köşe tanıma, istatistiksel örüntü tanıma, şablon eşleme gibi çeşitli yöntemler kullanılmaktadır [10, 13].

Bilgisayar kullanımı ile görüntü tanımının yaygınlaşmasıyla, tarım alanında, ürün kalitesinin gözlenmesi [14], ürün sulama [15], ilaçlama, hasat, ürün sınıflandırma, ürün gelişmelerinin gözlenmesi gibi çalışmalar yapılmaktadır [16]. Yine tarım alanında, görüntü işleme tekniklerinin kullanılması ile yapılan çeşitli çalışmalarda şeftali [17, 18], elma [18, 19], buğday [20], fındık [21], kiraz [22, 23], ceviz [24], badem [25] vb. meyveler sınıflandırılmakta ve özellikleri belirlenmektedir. Bu özellikler saptanırken sayısal görüntü analizi, sınıflama, kümeleme gibi yöntemler kullanılmakta ve incelenen nesnelere boyut, cins veya kalite bakımından sınıflandırılması yapılabilmektedir.

Araştırılan ortamda bulunan aynı nesnelere tespiti ve sınıflandırılmasıyla ilgili yapılan çalışmalarda 3 aşamalı bir yöntem uygulanmaktadır [10]. Bu aşamalar;

Aşama 1 – Görüntü Ön İşleme

Bu aşamada ilk olarak kameradan elde edilen görüntüye filtreleme işlemi uygulanmakta, böylece istenmeyen görüntülerin giderilmesi ve gereksiz detayların minimuma indirilmesi sağlanmaktadır. Daha sonra renkli olan görüntü grileştirilerek eşikleme işlemiyle yalnız ilgili nesnelere ait kısımlar elde edilmektedir. Bu işlemden sonra siyah ve beyaz görüntü oluşturulmakta ve siyah olan yerlerde beyaz, beyaz olan yerlerde ise siyah renkte istenmeyen noktalar bulunmaktadır. Oluşan ikili görüntüdeki istenmeyen görüntüleri silmek amacıyla morfolojik işlem uygulanmakta, ikili görüntüdeki beyaz bölgeleri daraltmak ve siyah bölgelerdeki beyazlıkları gidermek için aşındırma işlemi uygulanmaktadır. Beyaz bölgenin sınırlarını genişletmek ve bu bölgedeki siyahlıkları gidermek için ise genişleme işlemi uygulanmaktadır.

Aşama 2 – Nesne Bulma ve Özellik Çıkarımı İşlemi

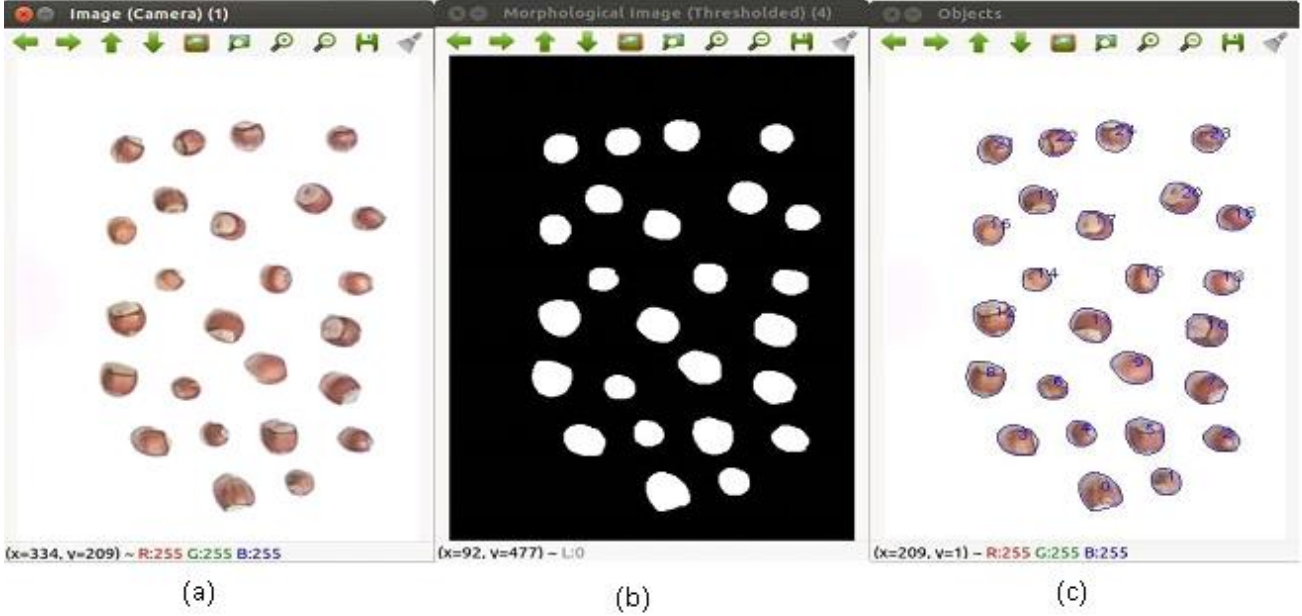
Bu aşamada görüntü, ön işleme aşamasından geçirilerek elde edilen ikili görüntü üzerinde nesnelere bulunması ve her bir nesneye ait özelliklerin çıkarımı işlemleri gerçekleştirilmektedir. Özellik çıkarım vektörlerinde nesnelere görüntü düzleminde kaplanmış olduğu alan, nesne boyları ve nesne merkezine ait koordinatlar yer almaktadır.

Aşama 3 – Sınıflandırma İşlemi

Sınıflandırma işlemiyle fiziksel veya soyut nesnelere benzer nesne sınıfları içerisinde gruplandırılmaktadır. Bu işlem nesnelere birbirleri ile benzer olup olmamasına göre gerçekleştirilmektedir. Sınıflandırma analizinde desen, nokta veya nesnelere doğal olarak gruplandırılması yapılmakta ve bu analizle çok

değişkenli özellikler içeren veriler kümelendirilmektedir. Kümeleme yöntemi görüntü tanıma, veri analizi, görüntü işleme, market araştırmaları vb. gibi çeşitli alanlarda kullanılmaktadır. Ortamdaki nesnelere genellikle; nesnelere alan, çap, yarıçap, genişlik, yükseklik gibi özellikleri esas alınarak sınıflandırılmaktadır.

Görüntü işleme teknikleri ve kümeleme yöntemleri kullanılarak fındık meyvesinin tespit ve sınıflandırılmasının yapıldığı bir çalışmada bu 3 aşamaya ait görüntüler Şekil 1'de gösterilmektedir [10].



Şekil 1. Görüntü işleme teknikleri ve kümeleme yöntemleri kullanılarak fındık meyvesinin tespit ve sınıflandırılmasının yapıldığı deneysel bir çalışmadan alınan örnek görüntü, (a) Kameradan alınan görüntü, (b) Ön işleme aşamasından sonra elde edilen görüntü, (c) Nesne bulma ve özellik çıkarım işleminde elde edilen görüntü [10].

Şekil 1 a'da kameradan alınan ilgili kısmın görüntüsü çalışma alanı dışında kalan kısım kesilerek verilmektedir. Şekil 1 b'deki görüntü, ön işleme aşamasında üzerinde filtreleme, grileştirme, eşikleme ve morfolojik işlem basamakları uygulandıktan sonra elde edilen görüntüdür. Bu görüntü nesne bulma ve özellik belirleme aşamasına girdi olarak verilmektedir. Şekil 1 c'de ise ortamdaki ilgilenilen nesnelere dış hatları ve indis numaraları sunulmakta, çalışmada kullanılacak alan, çap, yarıçap ve merkez noktasına ait koordinatlar elde edilmektedir [10].

GIDA ANALİZLERİNDE DENETİMLİ ÖRÜNTÜ TANIMA TEKNİKLERİ

Denetimli örüntü tanıma teknikleri, bilinen sınıflardan birinde ölçüm modeline dayanarak yeni bilinmeyen örnekleri sınıflandırmak için örneklerin sınıf üyeliği hakkındaki bilgileri belirli bir gruba (sınıf veya kategori) atamada kullanılmaktadır [1, 2]. Söz konusu teknikler, profil oluşturma, parmak izi, kimlik doğrulama, tahakkuk tespiti, gıda kalite değerlendirmesi, veri yorumlama gibi çeşitli amaçlarla çok çeşitli kimyasal verilere (kromatografik, spektrometrik, spektrofotometrik, spektroskopik, duyuusal vb.) uygulanabilmektedir [26].

Doğrusal Tanımlayıcı Analiz (Linear Discriminant Analysis, LDA)

LDA ayırım yapılacak sınıfların doğrusal olduğu bir çeşit tanımlayıcı analizdir. En çok kullanılan denetimli örüntü tanıma yöntemi olan LDA lineer tanımlama fonksiyonlarının belirlenmesinde kullanılır. Söz konusu yöntemde sınıflar arası varyans (veri setindeki değerlerin ortalamaya göre dağılımı ve değişimi) oranı en üst düzeyde olurken, sınıf içi varyans oranı en alt düzeyde olmaktadır. LDA'da sınıfların çok değişkenli normal dağılımı takip etmesi ve doğrusal olarak ayrılması gerekmektedir. LDA, temel bileşen analizinin (PCA) daha küçük boyutlu bir hiper düzlem belirlemesi anlamında bir özellik azaltma yöntemi olarak da düşünülmektedir. Bununla birlikte LDA, verilen sınıflar arasında maksimum ayrışmayı sağlayan bir yön seçmektedir. Düzenli tanımlayıcı fonksiyonun (RDA) alt durumları olan kuadratik tanımlayıcı fonksiyonu (QDA) ve Bayes sınıflandırma fonksiyonu gibi çeşitli başka yöntemler de kullanılabilir. QDA parabolik sınıflar oluşturmada, böylece uzaydaki nesnelere LDA'ya göre dağılımdaki kısıtlamalara daha az maruz kalmaktadır, ancak benzer şekilde örneklerin sayısının değişken sayısından daha yüksek olmasını gerektirmektedir. Daha fazla nesne gerektirmeden kısıtlamalara daha az maruz kalan RDA, LDA ve QDA'ya kıyasla daha avantajlıdır. Bayes yaklaşımı, her bir sınıfın üyeliğinin önceki bir olasılığa sahip olması prensibine dayanmaktadır. Kanonik değişken/korelasyon analizi

(CVA/CCA) de LDA'dan farklı olan diğer bir tanımlama tekniğidir ve LDA üyelik bilgisi içeren bir vektör kullanırken CVA bir matris kullanmaktadır [1, 2, 6, 27].

Kısmi En Küçük Kareler Tanımlayıcı Analizi (Partial Least Squares Discriminant Analysis, PLS-DA)

LDA tekniğinde, değişken sayısının nesnelere sayısını geçmemesi gerekliliği teknik açısından bir kısıtlamaya neden olmaktadır. Bu problemin giderilmesi çözümün reddedilmesi, veya kısmi en küçük kareler (PLS) gibi yöntemlerin uygulanması ile mümkün olmaktadır. Kısmi en küçük kareler modellemesi, çok değişkenli bir yöntem olup bu değişkenler giriş değişkenlerindeki ilgili varyasyonları olabildiğince açıklayan ve Y'deki hedef değer ile maksimum korelasyona sahip olan giriş matrisindeki bileşenleri (X) bulma ilkesine dayanmaktadır. PLS analizi X ve Y matrisleri arasındaki kovaryansı (iki değişkenin birbiriyle ilişkisi) en üst düzeye çıkararak değişkenleri ve yönleri bulmayı amaçlamaktadır. En uygun değişken sayısı, çapraz doğrulama veya harici test kümeleri kullanılarak tahmin edilmektedir. PLS-DA'da 1 ve 0'lar ile sahte bir Y matrisi oluşturulurken, X matrisi orijinal verilerden oluşmakta ve her sınıf için bir model geliştirilmektedir. Y'deki belirli bir sütunun ögesi 1'e ve diğer sütunların ögeleri 0'a ne kadar yakınsa o nesne belirli bir sınıfın üyesi olarak kabul edilmektedir. Sadece X'in bilgilerini kullanan PCA'nın aksine PLS'de X ve Y'deki bilgiler birlikte değerlendirilmekte, hatalar hesaba katılmakta ve eşit olarak dağıtıldığı varsayılmaktadır. Söz konusu yöntemin az objesi olan veri kümeleri için uygun olduğu ifade edilmektedir. PLS için her biri çalışılan duruma bağlı olarak belirli avantajları olan çeşitli algoritmalar mevcut bulunmaktadır. Bunlar arasında doğrusal olmayan yinelemeli kısmi en küçük kareler algoritması (NIPALS) temel bileşenlerin tek tek hesaplanmasını sağlamaktadır. PLS'de kullanılan algoritmalar hakkında daha fazla bilgi literatürde yer almaktadır [2, 6, 28, 29].

K-En Yakın Komşular (K-Nearest Neighbor, kNN)

kNN yöntemi bilinmeyen bir nesne ile eğitim setinin her bir nesnesi arasındaki mesafelerin belirlenmesine dayanmakta ve genellikle öklid mesafesi kullanılmaktadır. Ancak ilişkili olmayan değişkenler için korelasyon temelli yöntemler tercih edilmektedir. Nesnenin belli bir sınıfa üyeliği için eğitim seti olarak kabul edilen gruplara olan en düşük mesafe seçilmektedir. Bilinmeyen numuneye k-en yakın nesnelere seçilerek çoğunluk kuralı uygulanmakta ve böylece bilinmeyen nesne, k nesnesinin çoğunluğuna ait olduğu grupta sınıflandırılmaktadır. K seçimi farklı k değerleri ile tahmin yeteneğinin hesaplanmasıyla optimize edilerek sıklıkla 3 ve 5 gibi küçük k değerleri tercih edilmektedir. Yöntemin matematiksel basitliğinden dolayı sınıflandırma sonuçları diğer örüntü tanıma tekniklerinden daha iyi elde edilebilmektedir. Ayrıca değişkenlerin normal dağılımı gibi istatistiksel varsayımlardan bağımsız olması, etkinliğinin sınıfların uzay dağılımına bağlı olmaması gibi çeşitli avantajlar sunmaktadır. Söz konusu yöntemin LDA ile benzer sınıflamaları olmakla beraber eğer örnek sayısında

büyük farklılıklar var ise kNN iyi çalışmamaktadır. Çok sayıda örnek olduğunda hesaplamaların aşırı derecede yavaşlayabilme olasılığı bulunmaktadır. Bu nedenle sonuçlar grafiksel olarak ifade edilememektedir [6].

Sınıf Analogunun Esnek Bağımsız Modellemesi (Soft Independent Modelling of Class Analogy, SIMCA)

SIMCA, sınıf modelleme tekniklerinden en çok kullanılan yöntem olup her kategori bağımsız olarak PCA kullanılarak modellenmekte ve farklı sayıda ana bileşenlerle açıklanabilmektedir. Eğitim setindeki her bir sınıf için ana bileşenlerin sayısı çapraz doğrulama ile belirlenmektedir. Bu şekilde her bir sınıftaki varyasyonun çoğunu hesaba katarak yeterli sayıda temel bileşen korunurken, sınıf modelinde sözde ikincil veya gürültü yüklü ana bileşenler dahil edilmemiş olmaktadır. SIMCA ile sınıf mesafesi, modelleme ve ayırıcı güçler belirlenmektedir. Sınıf mesafesi ana bileşen modellerine olan geometrik uzaklık olarak hesaplanmaktadır. Ayırıcı güç ise bir değişkenin sınıfları ne kadar iyi ölçtüğünün göstergesidir. Bu yöntem ile elde edilen sonuçlar grafiksel olarak görselleştirilebildiğinden aykırı değerler, alt gruplamalar ve sınıf içi yapı hakkında bilgi sağlanmaktadır. Alan üzerinde sınıf 1, sınıf 2, sınıf 1 ve 2'nin üst üste gelmesi ve aykırı bölge (sınıf 1, 2 ve 1-2 kısımlarının dışında kalan) olmak üzere 4 bölge oluşturularak bir sınıflandırmanın ne kadar kesin olduğunun görselleştirilmesi mümkün olmaktadır [1, 2, 3, 6].

Çok Değişkenli Normal Dağılıma Dayalı Basit Sınıf Modelleme Tekniği (Unequal Dispersed Classes, UNEQ)

UNEQ, çok değişkenli normal dağılmış grupların varsayımına dayanan sınıf modelleme tekniğidir. Bu yöntem değişken sayısı düşük olduğunda uygulanabilmektedir. UNEQ kullanılarak bir sınıf içindeki nesnelere arasındaki benzerliklere dayanarak ayrı ayrı eğitim sınıflarının her biri için sınıflandırma fonksiyonları geliştirilmektedir. Bu nedenle eğitim sınıflarının her biri için örüntü alanında bir hiper düzlem tanımlanarak yeni bir nesnenin belirli bir sınıfa ait olup olmadığına karar vermek için bir aykırı test grubu kullanılmaktadır. UNEQ, sınıfın merkezindeki Mahalanobis mesafesine veya genelleştirilmiş mesafeye dayanmakta ve bu mesafe kritik bir değeri aştığında nesne aykırı olarak kabul edilmekte, yani sınıfın bir parçası olmamaktadır. Yöntem dengesiz veri kümelerine çok duyarlı olduğundan homojen popülasyon gerektirmektedir [2, 6, 30].

Sınıflandırma ve Regresyon Ağaçları (Classification and Regression Trees, CART)

CART, verilerin tekrar tekrar gruplara ayrıldığı bir ağaç oluşturma yöntemidir. CART'ın amacı en büyük olan X değerleri kümesinden bağımsız değişkenleri seçerek Y yanıtını açıklamaktır. Ağaç, dallara bağlı düğümlerle sonuçlanan özyinelemeli bir şekilde inşa edilmektedir. "İkili" terimi, bir karar ağacında bir "düğüm" ile temsil edilen her nesne grubunun yalnızca iki gruba

ayrılabilirliğini ifade etmektedir. Daha sonra iki yeni düğüme ayrılan bir düğüm, ana düğüm olarak adlandırılmakta ve iyi yeni düğümlere ise çocuk düğüm denilmektedir. Çocuk düğümleri olmayan düğümler ise terminal düğüm olarak adlandırılmaktadır. CART yöntemi temel olarak 3 adımdan oluşmaktadır. İlk olarak veri kümesi oluşturulmakta ve tüm örnekleri içeren bir kök düğümden başlanarak her ebeveyn düğüm en iyi ayırıcı tarafından iki çocuk düğüme ayrılmaktadır. Burada bölme işlemi tüm terminal düğümler bazı kriterleri karşılayana kadar devam etmektedir. İlk adımda elde edilen karar ağacı fazlalık verme eğiliminde olabildiğinden bilinmeyen örnekler için zayıf tahminler yapılabilmektedir. Bu problem ikinci adım olan budama işlemi ile çözülebilmekte ve bu işlem ile birlikte alt ağaçlar oluşturulmaktadır. Üçüncü aşamada ise, çapraz doğrulama gerçekleştirilerek veya harici bir test seti kullanılarak en iyi budanmış alt ağaç, yeni veriler için temel olarak seçilmektedir [6, 31, 32].

Destek Vektör Makinesi (Support Vector Machine, SVM)

SVM, istatistiksel öğrenme teorisine dayanan hem sınıflandırma hem de regresyon problemleriyle başa çıkmak için uygulanan denetimli bir öğrenme tekniğidir. SVM'nin amacı, her iki sınıfı tam olarak ayıran ve sadece eğitim setini değil, aynı zamanda bilinmeyen örnekleri de sınıflandıran her iki sınıfı da ayrı ayrı birleştirmektedir. En uygun sınır her iki kümeden en uzak hiper düzlem olarak yani bu kümeler arasındaki orta nokta olarak tanımlanmaktadır. Kümelerin dağılımı bilinmemekle birlikte, bu sınırın kümelerin en uygun sınıflandırması olması beklenmektedir, çünkü bu sınır her iki kümeden en izole olanıdır. Sınırlara en yakın eğitim vektörleri destek vektörleri olarak adlandırılmaktadır. Sınıflar doğrusal olmayan bir sınırla ayrıldığında, sınırı bulmak için çekirdek yöntemi kullanılmaktadır. Çekirdek yönteminin temel kavramı, vektör uzayının sınıfların doğrusal olarak ayrılabilirliği daha yüksek boyutlu bir alana dönüştürülmesinden oluşmaktadır ve regresyon durumunda küçük hatalar dikkate alınmadan doğrusal regresyon oluşturulmaktadır. [6, 33-35].

Yapay Sinir Ağları (Artificial Neural Networks, ANN)

ANN, büyük ölçüde veri işleme ve bilgi temsili için hesaplamalar yapabilen yapay nöronlar veya düğümler adı verilen birbirine bağlı, uyarlamalı basit işlem elemanlarından oluşan yapılar olarak tanımlanmaktadır. İleri beslemeli ANN paralel bir yapıya yerleştirilmiş çok sayıda yapay nörondan oluşmaktadır. Nöronlar, her bağımsız değişken (X) için bir nöron içeren giriş katmanında sıralanmakta ve verinin işlendiği bir veya daha fazla gizli katman ile her bağımlı değişken (Y) için bir nöron içeren çıkış katmanı bulunmaktadır. Giriş katmanından gelen veriler, ağlar yoluyla ağırlık (W) olarak adlandırılan bağlantı katsayıları ile ilişkili olan sinapslar ile yayılmaktadır. En popüler, çok yönlü ve yaygın olarak kullanılan ağ türü, geri yayımlı ANN'lerdir. Geri yayılım terimi, çıktı tarafında hesaplanan hatanın çıkış katmanından gizli katmana ve

son olarak da giriş katmanına doğru yayılma şekli ifade etmektedir. Geri yayılma ANN'lerinde nöronlar kısmen veya tamamen birbirine bağlanabilmektedirler. Geri yayılımın öğrenme stratejisi, bir önceki katmandan elde edilen hatayla orantılı olarak ağırlık bazında ölçümler yapılmasına dayanmaktadır. Bir ağın başlatılması, nöronlar arasındaki bağlantıların ağırlıklarına ve eşiklerine rastgele başlangıç değerlerinin atanmasını içermekte ve ağırlığın düzeltilmesi yinelemeli olarak gerçekleşmektedir. Eğitim aşamasında, öğrenme sürecine dahil olan tüm parametreler için değerler, net tahmin edilen çıktı ile doğru çıktı arasındaki hata hesaplanmaktadır. Gizli katmanların boyutu ve sayısı ile eğitim döngülerinin farklı değerleri test edilerek ve sonuç tahmininin doğruluğu kontrol edilerek değerlendirilmektedir. Bir ANN için optimal sayıda döngü, test setindeki hata minimum seviyeye ulaştığında elde edilmektedir. ANN'lerin büyük bir avantajı, girdi ve çıktı değişkenleri arasındaki ilişkinin nedensel bilgisinin gerekli olmamasıdır. Ayrıca, ANN'ler doğrusal olmayan yani aralarında nedensellik bulunmayan verilerin de değerlendirilmesini sağlama, ölçüm hatalarının varlığında doğru tahmin sağlayan gürültü duyarsızlığı, hızlı işleme ve donanım hatası toleransı anlamına gelen yüksek paralellik, değişen çevreye tepki olarak sistemin kendi yapısını güncellemesine (değiştirmesine) izin veren öğrenme ve uyarlanabilirlik, yapılan genelleme modelin kazanılmamış veriye uygulanmasını sağlaması gibi önemli ve pratik bilgi işleme özellikleri de sunmaktadır [36-39].

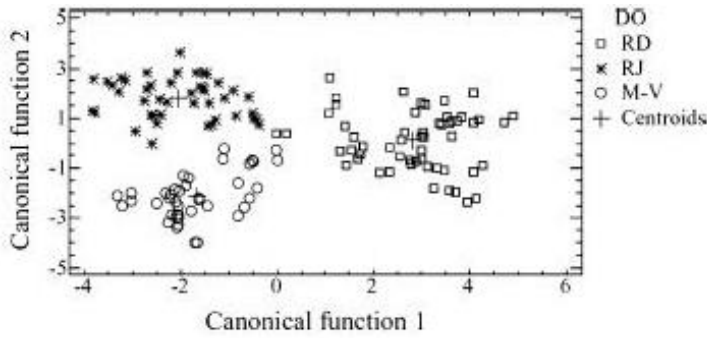
DENETİMLİ ÖRÜNTÜ TANIMA TEKNİKLERİNİN GIDA ANALİZLERİNDE UYGULANMASI İLE İLGİLİ ÖRNEKLER

Gıda analizlerinde kemometri (istatistik ve matematik ile birlikte bilgisayar kullanarak kimyasal verilerin işlenmesini kapsayan bir kimya disiplini) uygulamaları ile ilgili gıda ürünlerine genel bakış [40], balık [41], et [42], şarap [43, 44], bira [45] veya bal [46] gibi belirli gıdalar üzerine yapılan çalışmalar bulunmaktadır. Bunların dışında örüntü tanıma teknikleriyle ilgili olarak elektronik burun veya dil [47, 48], süt ürünlerinde doku tayini [49], şarap aroması [50], orta kızılötesi spektrometri (MIR) kullanılarak gıda doğrulaması [51], duyu analizi [52], bilgisayar görüntüsü ile gıda kalitesinin değerlendirilmesi [53], et bozulmalarının tespiti [54], çoklu element/izotop analizi [55] ve çeşitli gıdaların coğrafi kökenlerinin belirlenmesi üzerine literatürde çalışmalar bulunmaktadır [6].

Berrueta ve ark. [6] tarafından 2004-2006 yılları arasında bibliyografide bulunan gıda analizlerinde denetimli örüntü tanıma uygulamalarının ayrıntıları; girdi verileri, veri ön işleme yöntemleri, değişken seçimleri, denetimsiz/denetimli örüntü tanıma teknikleri ve sınıflandırma kriterleri şeklinde incelenmiştir. Şarap analizlerinde veri girişi olarak e-burun, e-dil, kimyasal veriler, renk bileşenleri, antosiyanin kompozisyonu gibi değişkenler seçilerek çeşitli denetimli örüntü tanıma teknikleri uygulanarak şaraplar coğrafi orijin, üzüm kültürü, depolama zamanı, yapım tekniği gibi kriterlere göre sınıflandırılmıştır. Yenilebilir yağ analizlerinde yağ

asit içeriği, duyu verileri, spektra verileri kullanılarak yağlar botanik orijin, coğrafi orijin, kaliteli veya kusurlu şeklinde kriterlere göre sınıflandırılmıştır. Bal analizlerinde nem, pH, iletkenlik, HMF içeriği, renk bileşenleri verileri kullanılarak bal ürünlerinde genelde çiçek kökenine göre sınıflandırılma yapılmıştır. Süt ürünleri analizlerinde veri girişi olarak duyu verileri, peptit ve aminoasit profilleri seçilerek ürünler marka, ürün çeşidi, olgunlaşma zamanı gibi kriterlere göre sınıflandırılmıştır. Et analizlerinde protein içerikleri, multispektral görüntüler, dioksin ve dibenzofuranlar veri girişi olarak seçilerek et ürünleri dondurulmuş saklama zamanı, hayvan kökeni ve asit oranı şeklinde sınıflandırılmıştır. Meyve ürünleri analizlerinde ise polifenol, prosiyanidin, e-burun giriş verileri kullanılarak meyveler çeşit, coğrafi orijin, olgunluk durumu, kalite gibi kriterlere göre sınıflandırılmıştır [6]. Tüm bu denetimli örüntü tanıma teknikleri ile ilgili yapılan çalışmalar incelendiğinde gıda örneklerini farklı kriterlere göre sınıflandırmak için çeşitli modellerin kullanılabilirliği görülmektedir. Bu modeller kullanılarak coğrafi köken, hayvansal köken, botanik köken, teknolojik süreç, kalite durumu, taşıma tespiti gibi kategorilerde ayrımların yapılabileceği çeşitli sınıflar oluşturulmaktadır.

P'erez-Magari'no ve ark. [56] 70 örnekten oluşan gül şaraplarını coğrafi kökenlerine göre sınıflandırmışlardır.

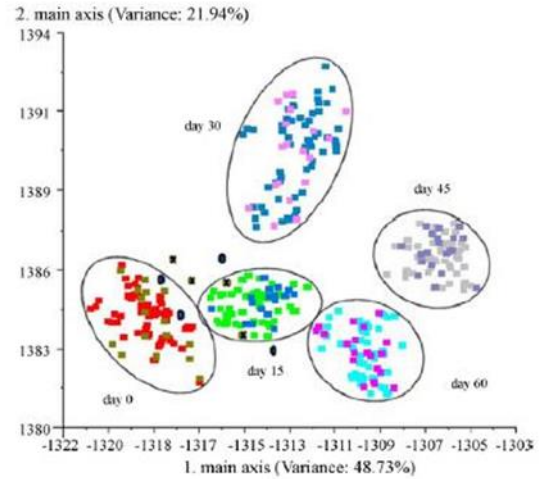


Şekil 2. Coğrafi kökenli (RD, Ribera del Duero; RJ, Rioja; M-V, La Mancha ve Valdepeñas) gül şaraplarının sınıflandırılması için oluşturulan kanonik değişkenler (kökler) çizimi [56]

Cozzolino ve ark. [58] tarafından ham ve ikinci türev vis-NIR (200-2500 nm) spektral verileri kullanılarak iki botanik kökene ait beyaz şaraplar (Riesling çeşidinin 144 numunesi ve Chardonnay çeşidinin 125 numunesi) arasında ayırım yapılmıştır. PCA kullanılarak veri boyutunun azaltılması gerçekleştirilmiştir. En iyi sınıflandırma yetenekleri, 400 ve 1100 nm arasındaki spektrallerin ikinci türev verileri kullanılarak gerçekleştirilmiştir (çapraz doğrulama ile isabetlerin %100'ü). Şekil 4'te Riesling ve Chardonnay şarapları için doğru tahminlerin sırasıyla %100 ve %98,4'ünü

Başlangıçtaki 19 başlangıç girişi değişkeninden (Atomik Absorpsiyon Spektrometresi ile 7 metal, geleneksel kimyasal yöntemlerle 6 polifenolik bileşik ailesinin toplam içeriği, 2 renk parametresi ve 4 klasik oenolojik parametre), sadece 10 değişken LDA ile seçilmiştir. Çalışmada 4 farklı köken belirlenmiştir (Ribera del Duero (RD), Rioja (RJ), La Mancha (M) ve Valdepeñas (V)). Valdepeñas ve La Mancha'dan alınan örnekler bir sınıf (M-V) olarak kabul edilmiştir. Bu bölgelerin coğrafi olarak yakın olması nedeniyle benzer iklimoloji ve detaylandırma uygulamaları göstermektedir. Şekil 2'de, numunelerin ilk iki kanonik varyasyon tarafından tanımlanan alan üzerindeki grafiği görülmektedir. LDA modelinin tanıma yetenekleri RD için %98.2 ve RJ ile M-V için %100 iken tahmin yetenekleri RD için %95.3, RJ için %98.1 ve M-V için %98.0 olarak belirtilmiştir [56].

Hernandez-Gomez ve ark. [57] 10 sensör ile donatılmış bir e-burun kullanarak mandalinaların farklı olgunluk durumlarını (400 örnek) karakterize etmişlerdir. Bu amaçla, mandalinaları beş farklı toplama tarihinde (her 15 günde bir) hasat etmişlerdir. LDA, bir dış doğrulama kullanılarak gerçekleştirilmiş ve sırasıyla %98 ve %92'lik küresel tanıma ve öngörü yeteneklerine ulaşılmıştır. Şekil 3'te ilk iki kanonik varyasyonun çizimi, 0 ve 15 günlüklerden sadece bazı örneklerin yanlış sınıflandırıldığını göstermektedir [57].

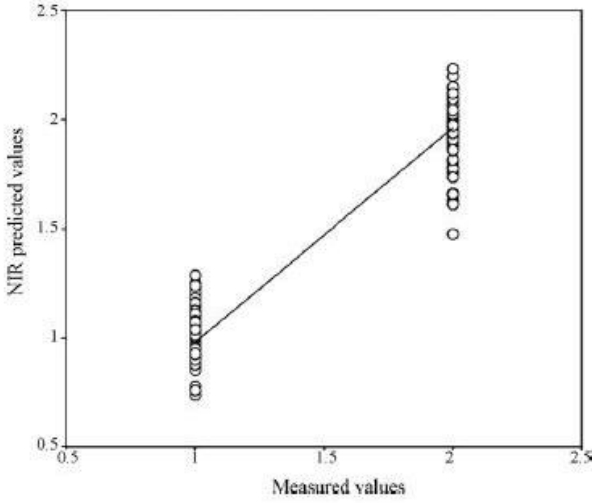


Şekil 3. E-burun verileri kullanılarak mandalina farklı olgunluk durumlarının sınıflandırılması için hazırlanan LDA grafiği. Eğitim seti: kırmızı, yeşil, açık mavi, turkuaz mavi ve gri kareler. Harici test seti: koyu yeşil, koyu mavi, pembe, fuşya, mor [57].

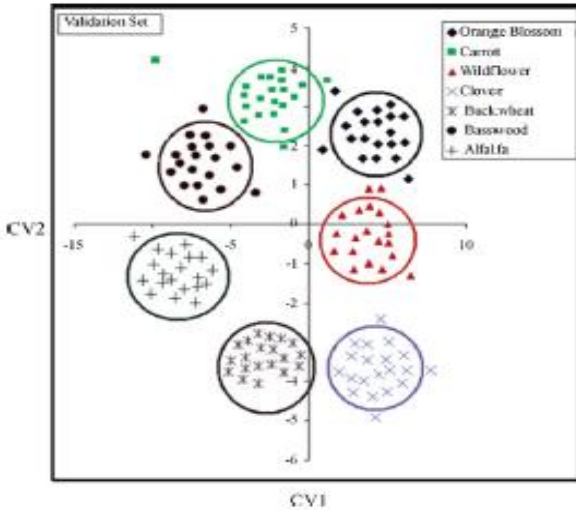
sağlayan, modelin harici doğrulaması ile ilgili sonuçlar görülmektedir [58].

Tewari ve Irudayaraj [59], FT-IR ve z-burnu ile yedi farklı botanik kökene (yonca, karabuğday, ıhlamur, kır çiçeği, portakal çiçeği, havuç ve yonca) ait balları (350 örnek) karakterize etmişlerdir. Veriler SD'ye bölünmesinden sonra normalleştirilmiş ve daha sonra bir PCA değişken indirgemesi gerçekleştirilmiştir. İlk altı veya yedi PC'nin skorları, CVA için giriş verileri olarak seçilmiştir. Şekil 5'te elde edilen tanıma ve tahmin yeteneklerinin

sırasıyla %97 ve %96'dan (yedi köken ortalamaları) daha yüksek olarak elde edildiği görülmektedir [59].



Şekil 4. İki farklı üzüm çeşidinden elde edilen beyaz şaraplar arasındaki ayrım grafiği: PLS-DA (sınıf 1, Riesling şarapları ve sınıf 2, Chardonnay şarapları) tarafından elde edilen validasyon setine yönelik tahmin edilen değerlere karşı ölçülen değerler [58]



Şekil 5. Farklı çiçeklerden elde edilen bal örneklerinin sınıflandırılması için ilk iki kanonik varyasyonun (CV1 ve CV2) CVA grafiği [59]

Bu çalışma örnekleri incelendiğinde, denetimli örüntü tanıma tekniklerinin doğru bir şekilde kullanıldığında faydalı sonuçlar elde etmeye olanak sağlayan yöntemler olduğu görülmektedir. Bu anlamda, her bir kategorideki değişkenliği doğru temsil etmeyen az sayıda örnek kullanılması gibi yetersiz uygulamalardan kaçınmaya özel önem verilmelidir. Dengesiz örnek kümelerinin kullanılması ve modellerin doğrulamasını yapmak ta önem taşımaktadır [6].

SONUÇ

Analitik kimyada, modern analitik araçların sağladığı yüksek miktarda analitik bilgi nedeniyle veri analizi uygulamaları temel bir basamak haline gelmiştir.

Denetimli örüntü tanıma, ölçülen özelliklerin modeline dayanarak, önceden tanımlanmış bir örnek sınıfına bilinmeyen örnekler atamak amacıyla deneysel verilere dayanan bir sınıflandırma modeli oluşturmayı amaçlamaktadır. Gıda analizlerinde denetimli örüntü tanıma kullanımı, bu alanın birçok yönünü kapsayarak ve kullanım oranı gün geçtikçe katlanarak artmaktadır. Mevcut çalışmalar ve araştırmalar gözden geçirildiğinde, çok değişkenli tekniklerin istatistiksel temellerinin anlaşılmasının hala gerekli olduğu sonucuna varılabilir. Denetimli örüntü tanıma teknikleriyle incelenen gıdalara örnek olarak şarap, yenilebilir yağ, bal, süt ürünleri, et, meyveler, içecekler, tahıllar, balık, katkı maddeleri vb. ürünler verilebilir. Özetlenen örnek çalışmalar, gıda kimyasında denetimli örüntü tanıma yöntemlerinin potansiyel yararlılığını ve çok çeşitli yöntemlerin rutin olarak uygulandığını göstermektedir. Ana bileşen analizinin, herhangi bir denetimli örüntü tanıma tekniğini kullanılmadan önce yapılması denetimsiz örüntü tanıma uygulamasının temelini oluşturmaktadır ve yapılan çalışmalarda bu uygulamanın yaygın olarak kullanıldığı görülmektedir. LDA en sık kullanılan denetimli örüntü tanıma tekniğidir. Elde edilebilecek iyi sonuçları sunmasına rağmen ise QDA ve CART nadiren kullanıldığı saptanmıştır.

Yapılan çalışmalar incelendiğinde, dikkate alınan sınıfların geçerliliği, araştırmacıların temsili bakış açısı ve belirli noktalara aşırı takılmaları nedeniyle eksiklikler olduğu görülmektedir. Çalışmaların amacına ulaşması, gıda analizlerinde en çok kullanılan denetimli örüntü tanıma tekniklerinin temellerinin gözden geçirilerek, ölçülen verilerin pratik gereksinimlerine özel bir vurgu yapılması ile söz konusu olabilecektir. Bu uygulamanın, gıda analizleri konusunda hızlı, pratik ve güvenilir sonuç olanağı vermesinden dolayı önümüzdeki yıllarda kullanım olarak büyük bir potansiyele sahip olduğu düşünülmektedir.

KAYNAKLAR

- [1] Lavine, B.K. (2000). Encyclopedia of Analytical Chemistry. John Wiley & Sons Ltd., Chichester, New York.
- [2] Massart, D.L., Vandeginste, B.G.M., Buydens, L.M.C., De Jong, S., Lewi, P.J., Smeyers-Verbeke, J. (1997). Handbook of Chemometrics and Qualimetrics: Part A, Elsevier, Amsterdam, 207p.
- [3] Brereton, R.G. (2003). Chemometrics: Data Analysis for the Laboratory and Chemical Plant. Wiley, Chichester, 119p.
- [4] Mazzatorta, P., Benfenati, E., Lorenzini, P., Vighi, M. (2004). QSAR in ecotoxicity: an overview of modern classification techniques. *Journal of Chemical Information and Computer Sciences*, 44(1), 105-112.
- [5] Todeschini, R., Ballabio, D., Consonni, V., Mauri, A. Pavan, M. (2007). CAIMAN (Classification And Influence Matrix Analysis): A new approach to the classification based on leverage-scaled functions. *Chemometrics and Intelligent Laboratory Systems*, 87(1), 3-17.
- [6] Berrueta, L.A., Alonso-Salces, R.M., Héberger, K. (2007). Supervised pattern recognition in food

- analysis. *Journal of Chromatography A*, 1158(1-2), 196-214.
- [7] Koyuncu, İ. (2016). İleri Örüntü Tanıma Teknikleri ve Uygulamaları. <http://docplayer.biz.tr/3182643-İleri-oruntu-tanima-teknikleri-ve-uygulamaları-icerik.html>.
- [8] Anonim (2014). Örüntü Tanıma. http://ehm.kocaeli.edu.tr/dersnotlari_data/kgullu/Oruntu%20Tanima/Sunu1_2.pdf.
- [9] Samtaş, G., Gülesin, M. (2011). Sayısal Görüntü İşleme ve Farklı Alanlardaki Uygulamaları. *Electronic Journal of Vocational Colleges*, 2(1), 85-97.
- [10] Solak, S., Altınışik, S. (2018). Görüntü işleme teknikleri ve kümeleme yöntemleri kullanılarak fındık meyvesinin tespit ve sınıflandırılması. *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 22(1), 56-65.
- [11] Viola, P., Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (Vol. 1). IEEE.
- [12] Hussin, R., Juhari, M.R., Kang, N.W., Ismail, R.C., Kamarudin, A. (2012). Digital image processing techniques for object detection from complex background image. *Procedia Engineering*, 41, 340-344.
- [13] Sonka, M., Hlavac, V., Boyle, R. (2014). *Image Processing, Analysis, and Machine Vision*. Cengage Learning, Stamford, USA, 37p.
- [14] Wu, D., Sun, D.W. (2013). Colour measurements by computer vision for food quality control—A review. *Trends in Food Science & Technology*, 29(1), 5-20.
- [15] Hof, A., Wolf, N. (2014). Estimating potential outdoor water consumption in private urban landscapes by coupling high-resolution image analysis, irrigation water needs and evaporation estimation in Spain. *Landscape and Urban Planning*, 123, 61-72.
- [16] Latha, M., Poojith, A., Reddy, B.A., Kumar, G.V. (2014). Image processing in agriculture. *International Journal of Innovative Research In Electrical, Electronics, Instrumentation and Control Engineering*, 2(6), 1562-1565.
- [17] Kurtuluş, F., Vardar, A., Kavdır, İ. (2013). Bahçe Koşullarında Alınmış Renkli Görüntülerde Doku ve Şekil Özellikleriyle Genç Şeftali Meyvelerinin Saptanması. *Tarım Makinaları Bilimi Dergisi*, 9(2), 141-148.
- [18] Sert, E., Taşkın, D., Suçsuz, N. (2010). Görüntü İşleme teknikleri ile şeftali ve elma sınıflandırma. *Trakya Üniversitesi Fen Bilimleri Dergisi*, 11(2), 82-88.
- [19] Sofu, M., Er, O., Kayacan, M.C., Çetişli, B. (2013). Elmaların görüntü işleme yöntemi ile sınıflandırılması ve leke tespiti. *Gıda Teknolojileri Elektronik Dergisi*, 8(1), 12-25.
- [20] Demirbaş, H.Y., Dursun, İ. (2007). Buğday tanelerinin bazı fiziksel özelliklerinin görüntü işleme tekniğiyle belirlenmesi. *Ankara Üniversitesi Ziraat Fakültesi Tarım Bilimleri Dergisi*, 13(3), 176-185.
- [21] Bayrakdar, S., Çomak, B., Başol, D., Yücedağ, İ. (2015). Determination of type and quality of hazelnut using image processing techniques. In *Signal Processing and Communications Applications Conference (SIU)*, May 2015, 616-619p.
- [22] Balcı, M., Altun, A.A., Taşdemir, Ş. (2016). Görüntü işleme teknikleri kullanılarak Napolyon tipi kirazların sınıflandırılması. *Selçuk-Teknik Dergisi*, 15(3), 221-237.
- [23] Beyer, M., Hahn, R., Peschel, S., Harz, M., Knoche, M. (2002). Analysing fruit shape in sweet cherry (*Prunus avium* L.). *Scientia Horticulturae*, 96(1), 139-150.
- [24] Ercisli, S., Sayinci, B., Kara, M., Yıldız, C., Ozturk, I. (2012). Determination of size and shape features of walnut (*Juglans regia* L.) cultivars using image processing. *Scientia horticulturae*, 133, 47-55.
- [25] Antonucci, F., Costa, C., Pallottino, F., Paglia, G., Rimatori, V., De Giorgio, D., Menessati, P. (2012). Quantitative method for shape description of almond cultivars (*Prunus amygdalus* Batsch). *Food and bioprocess technology*, 5(2), 768-785.
- [26] Lavine, B.K. (2006). Pattern recognition. *Critical Reviews in Analytical Chemistry*, 36, 153-161.
- [27] Héberger, K., Csomós, E., Simon-Sarkadi, L. (2003). Principal component and linear discriminant analyses of free amino acids and biogenic amines in hungarian wines. *Journal of Agricultural and Food Chemistry*, 51(27), 8055-8060.
- [28] Hörchner, U., Kalivas, J.H. (1995). Simulated annealing type optimization algorithms: fundamentals and wavelength selection applications. *Journal of Chemometrics*, 9, 283-308.
- [29] Rogers, D., Hopfinger, A.J. (1994). Application of genetic function approximation to quantitative structure-activity relationships and quantitative structure-property relationships. *Journal of Chemical Information and Computer Sciences*, 34(4), 854-866.
- [30] Derde, M.P., Massart, D.L. (1986). UNEQ: a disjoint modelling technique for pattern recognition based on normal distribution. *Analytica Chimica Acta*, 184, 33-51.
- [31] Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984). *Classification and Regression Trees*. Wadsworth International Group, Belmont, CA, 131-156p.
- [32] Zhang, M.H., Xu, Q.S., Daeyaert, F., Lewi, P.J., Massart, D.L. (2005). Application of boosting to classification problems in chemometrics. *Analytica Chimica Acta*, 544, 167-176.
- [33] Burges, C.J.C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 121-167.
- [34] Cortes, C., Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20, 273-297.
- [35] Xu, Y., Zomer, S., Brereton, R.G. (2006). Support Vector Machines: A Recent Method for Classification in Chemometrics. *Critical Reviews in Analytical Chemistry*, 36(3-4), 177-188.
- [36] Barile, D., Coisson, J.D., Arlorio, M., Rinaldi, M. (2006). Identification of production area of

- Ossolano Italian cheese with chemometric complex approach. *Food Control*, 17(3), 197-206.
- [37] Basheer, I.A., Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 43(1), 3-31.
- [38] Gonçalves, E.C., Minim, L.A., Coimbra, J.S.R., Minim, VPR. (2005). Modeling sterilization process of canned foods using artificial neural networks. *Chemical Engineering and Processing: Process Intensification*, 44(12), 1269-1276.
- [39] Jain, A.K., Mao, J., Mohiuddin, K.M. (1996). Artificial neural networks: A tutorial. *Computer*, 29(3), 31-44.
- [40] Tzouros, N.E., Arvanitoyannis, I.S. (2001). Agricultural produces: synopsis of employed quality control methods for the authentication of foods and application of chemometrics for the classification of foods according to their variety or geographical origin. *Critical Reviews in Food Science and Nutrition*, 41(4), 287-319.
- [41] Arvanitoyannis, I.S., Tsitsika, E.V., Panagiotaki, P. (2005). Implementation of quality control methods (physicochemical, microbiological and sensory) in conjunction with multivariate analysis towards fish authenticity. *International Journal of Food Science and Technology*, 40, 237-263.
- [42] Arvanitoyannis, I.S., van Houwelingen-Koukaliaroglou, M. (2003). Implementation of chemometrics for quality control and authentication of meat and meat products. *Critical Reviews in Food Science and Nutrition*, 43(2), 173-218.
- [43] Arvanitoyannis, I.S., Katsota, M.N., Psarra, E.P., Soufleros, E.H., Kallithraka, S. (1999). Application of quality control methods for assessing wine authenticity: Use of multivariate analysis (chemometrics). *Trends in Food Science & Technology*, 10(10), 321-336.
- [44] Gishen, M., Damberg, R.G., Cozzolino, D. (2005). Grape and wine analysis - enhancing the power of spectroscopy with chemometrics. *Australian Journal of Grape and Wine Research*, 11(3), 296-305.
- [45] Siebert, K.J. (2001). Chemometrics in Brewing-A Review. *Journal of the American Society of Brewing Chemists*, 59(4), 147-156.
- [46] Arvanitoyannis, I.S., Chalhoub, C., Gotsiou, P., Lydakis-Simantiris, N., Kefalas, P. (2005). Novel quality control methods in conjunction with chemometrics (multivariate analysis) for detecting honey authenticity. *Critical Reviews in Food Science and Nutrition*, 45(3), 193-203.
- [47] Ampuero, S., Bosset, J.O. (2003). The electronic nose applied to dairy products: a review. *Sensors and Actuators B: Chemical*, 94(1), 1-12.
- [48] Vlasov, Y., Legin, A., Rudnitskaya, A., di Natale, C., Amico, A.D. (2005). Nonspecific sensor arrays (electronic tongue) for chemical analysis of liquids. *Pure and Applied Chemistry*, 77(11), 1965-1983.
- [49] Karoui, R., Mazerolles, G., Dufour, É. (2003). Spectroscopic techniques coupled with chemometric tools for structure and texture determinations in dairy products. *International Dairy Journal*, 13(8), 607-620.
- [50] Noble, A.C., Ebeler, S.E. (2002). Use of multivariate statistics in understanding wine flavor. *Food Reviews International*, 18(1), 1-20.
- [51] Downey, G. (1998). Food and food ingredient authentication by mid-infrared spectroscopy and chemometrics. *Trends in Analytical Chemistry*, 17(7), 418-424.
- [52] Sundberg, R. (2000). Aspects of statistical regression in sensometrics. *Food Quality and Preference*, 11, 17-26.
- [53] Du, J.C., Sun, D.W. (2006). Learning techniques used in computer vision for food quality evaluation: a review. *Journal of Food Engineering*, 72(1), 39-55.
- [54] Ellis, D.I., Goodacre, R. (2001). Rapid and quantitative detection of the microbial spoilage of muscle foods: current status and future trends. *Trends in Food Science & Technology*, 12(11), 414-424.
- [55] Kelly, S., Heaton, K., Hoogewerff, J. (2005). Tracing the geographical origin of food: The application of multi-element and multi-isotope analysis. *Trends in Food Science & Technology*, 16(12), 555-567.
- [56] Perez-Magarino, S., Ortega-Heras, M., Gonzalez-San Jose, M.L., Boger, Z. (2004). Comparative study of artificial neural network and multivariate methods to classify Spanish DO rose wines. *Talanta*, 62(5), 983-990.
- [57] Gómez, A.H., Wang, J., Hu, G., Pereira, A.G. (2006). Electronic nose technique potential monitoring mandarin maturity. *Sensors and Actuators B: Chemical*, 113(1), 347-353.
- [58] Cozzolino, D., Smyth, H.E., Gishen, M. (2003). Feasibility study on the use of visible and near-infrared spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins. *Journal of Agricultural and Food Chemistry*, 51(26), 7703-7708.
- [59] Tewari, J.C., Irudayaraj, J.M.K. (2005). Floral classification of honey using mid-infrared spectroscopy and surface acoustic wave based z-nose sensor. *Journal of Agricultural and Food Chemistry*, 53(18), 6955-6966.