T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

NÖTROZOFİK MANTIĞIN GÖRÜNTÜ BÖLÜTLEME VE ÖRÜNTÜ TANIMA ALANLARINDAKİ YENİ UYGULAMALARI

Yaman AKBULUT

Doktora Tezi
Anabilim Dalı: Elektrik-Elektronik Mühendisliği Teknolojileri
Danışman: Prof. Dr. Abdulkadir ŞENGÜR

TEMMUZ-2018

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

NÖTROZOFİK MANTIĞIN GÖRÜNTÜ BÖLÜTLEME VE ÖRÜNTÜ TANIMA ALANLARINDAKİ YENİ UYGULAMALARI

DOKTORA TEZİ

Yaman AKBULUT

(141138202)

Tezin Enstitüye Verildiği Tarih : 25.06.2018
Tezin Savunulduğu Tarih : 19.07.2018

Tez Danışmanı : Prof. Dr. Abdulkadir ŞENGÜR (F.Ü.)
Diğer Jüri Üyeleri : Prof. Dr. Hanifi GÜLDEMİR (F.Ü.)
Doç. Dr. Erkan DENİZ (F.Ü.)
Doç. Dr. M. Fatih TALU (İ.Ü.)
Dr. Öğr. Üyesi Ömer F. ALÇİN (B.Ü.)

TEMMUZ-2018

ÖNSÖZ

Bu tez çalışması Fırat Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Teknolojileri doktora programında hazırlanmıştır. Görüntü bölütleme ve örüntü tanıma alanlarında literatürde sayısız çalışmalar yapılmıştır ve yapılmaya devam etmektedir. Bu alanlara bir de nötrozofik mantık açısından bakmak, yorumlamak, yeni yöntemler eklemek ve mevcut olanları geliştirmek için bu çalışmaya başladık.

Bu süreçte tecrübesiyle bana yol gösteren, bilimsel katkılarıyla yardımcı olan, elindeki tüm imkanları sunmaktan çekinmeyen, sonuna kadar kullandıran, her türlü sorunda küçük bir dokunuşla çözüme ulaştıran ve bana bu çalışmayı yapma fırsatı veren danışman hocam Prof. Dr. Abdulkadir ŞENGÜR'e sonsuz teşekkürlerimi sunarım.

Enformatik bölüm başkanı Doç. Dr. Galip AYDIN'a ve bölümdeki çalışma arkadaşlarıma verdikleri destek için teşekkür ederim. Dr. Öğr. Üyesi Muzaffer ASLAN'a ve Dr. Öğr. Üyesi Ömer Faruk ALÇİN'e tez taslağının okunması, değerlendirilmesi ve düzenlemesi konularında katkılarından dolayı teşekkür ederim. Dr. Öğr. Üyesi Fatih ERTAM'a sağladığı lojistik destek ve ileriye dönük düşünceler için teşekkür ederim. Kaynakça stili için destek veren Arş. Gör. Ferhat UÇAR'a teşekkür ederim.

Illinois Springfield Üniversitesi'ndeki çalışmalarım süresince sağladığı destek için TÜBİTAK-BİDEB'e teşekkür ederim. Illinois Springfield Üniversitesi'ndeki danışmanım Assistant Profesor Yanhui GUO'ya Amerika Birleşik Devletleri'ne geliş sürecindeki yardımları ve araştırma projesindeki yönlendirmeleri için teşekkür ederim.

Tübitak bursu dolayısıyla yurt dışı görevlendirmesini çok kısa bir sürede gerçekleştiren başta Sayın rektörümüz Prof. Dr. Kutbeddin DEMİRDAĞ ve rektör yardımcısı Prof. Dr. Halil HASAR olmak üzere tüm yönetim kurulu üyelerine teşekkür ederim.

Ayrıca bütün zorluklara rağmen beni yalnız bırakmayan, beni sürekli destekleyen ve yaşam kaynağım olan çocuklarım Nazenin, Yiğit Hamza ve sevgili eşim Nevin AKBULUT'a en içten teşekkürlerimi sunarım.

Bu doktora tezinin nötrozofik mantık, görüntü bölütleme ve örüntü tanıma alanlarında çalışma yapacak tüm araştırmacılara küçük de olsa bir katkı sağlamasını dilerim. Saygılarımla.

Yaman AKBULUT

ELAZIĞ–2018

İÇİNDEKİLER

ÖZET

Latince'den tarafsız anlamında "neuter" ve Yunanca'dan beceri/bilgelik anlamında "sophia" kelimelerinden oluşan "Neutrosophy" Nötrozofi, 1980 yılında Smarandache tarafından tanıtılmıştır. Nötrozofi, felsefenin, mantığın, küme teorisinin, olasılık ve istatistik bilgisinin bir araya getirildiği bir felsefe dalıdır. <A> bir varlık ya da olay olarak tanımlansın; bu durumda <Non-A>, <A> olmayanlar ve <Anti-A> ise <A>'nın tersidir. Nötrozofi, belirsizliği temsil eden <Neut-A> adlı yeni bir kavram sunmaktadır. Ayrıca <Neut-A>, ne <A> ne de <Anti-A> olarak tanımlanamaz. Nötrozofi, t'nin doğru, i'nin belirsiz ve f'nin yanlış olduğu önermesine dayanan bulanık mantığın bir genelleştirilmesidir. t, i ve f; T, I ve F aralıklarından gerçek değerlerdir ve bunlar üzerinde hiçbir kısıtlama yoktur. Nötrozofi, belirsizlik bilgisine sahip olduğundan bulanık mantık ile çözülemeyen birçok problem nötrozofik mantık ile çözülebilmektedir.

Bu tez çalışmasında nötrozofik mantık kullanarak görüntü bölütleme ve örüntü tanıma alanlarında yeni yöntemlerin eklenmesi ve var olanların geliştirilmesi amaçlanmıştır. Önerilen algoritmalar ve yöntemler aşağıda belirtilmiştir:

- Gürültü ve aykırı veriye sahip veri kümelerinin kümelenmesi için Çekirdek Nötrozofik C-Ortalamalar (ÇNCO) kümeleme yöntemi,
- Dengesiz veri kümelerinin sınıflandırılması için yeni bir Nötrozofik Ağırlıklandırılmış Aşırı Öğrenme Makinesi (NAAÖM) yöntemi,
- Nötrozofik Küme tabanlı *k*-En Yakın Komşuluk (NK-*k*-EYK) sınıflandırıcısı,
- Etkili bir görüntü bölütleme algoritması için Nötrozofik Çizge Kesim (NÇK) önerilmiştir.

Önerilen bu algoritma ve yöntemlerin başarımını değerlendirmek amacıyla gerçek veri kümeleri, yapay veri kümeleri ve çeşitli görüntüler üzerinde deneysel çalışmalar yapılmıştır. Deneysel çalışmalar, önerilen yöntemlerin gürültü ve aykırı veri noktaları gibi belirsizlik içeren veri kümeleri ve görüntüler üzerinde dayanıklı ve etkin bir yapıya sahip olduğunu göstermektedir.

Bu tez çalışması, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK), Bilim İnsanı Destekleme Daire Başkanlığı (BİDEB), 2214/A Doktora Sırası Araştırma Burs Programı kapsamında desteklenmiştir.

Anahtar Kelimeler: Nötrozofi, nötrozofik mantık, nötrozofik küme, görüntü bölütleme, örüntü tanıma, nötrozofik c-ortalamalar, aşırı öğrenme makinesi, k-en yakın komşuluk.

SUMMARY

NOVEL APPLICATIONS OF NEUTROSOPHIC LOGIC IN THE FIELDS OF IMAGE SEGMENTATION AND PATTERN RECOGNITION

Neutrosophy, composed of “neuter” in neutral sense from Latin and “sophia” in the sense of skill/wisdom from Greek, was introduced by Smarandache in 1980. Neutrosophy is a philosophy that brings together the knowledge of philosophy, logic, set theory, probability, and statistics. <A> defined as an entity or event; in this case <Non-A> is non <A> and <Anti-A> is the inverse of <A>. Neutrosophy presents a new concept, <Neut-A>, which represents uncertainty. Furthermore, <Neut-A> cannot be defined as neither <A> nor <Anti-A>. Neutrosophy is a generalization of fuzzy logic based on the assumption that t is true, i is indeterminate, and f is false. t, i, f are in a range of T, I, F and they are real values and there are no restrictions on them. Since neutrosophy has knowledge of uncertainty, many problems that cannot be solved by fuzzy logic can be solved by neutrosophic logic.

In this thesis study, it is aimed to add new methods in image segmentation and pattern recognition areas using neutrosophic logic and to develop existing ones. Proposed algorithms and methods are as follows:

- Kernel Neutrosophic C-Means (KNCM) clustering method for clustering datasets with noise and outliers,
- A novel Neutrosophic Weighted Extreme Learning Machine (NWELM) method for classifying unbalanced data clusters,
- Neutrosophic Set-based k-Nearest Neighbor (NS-k-NN) classifier,
- Neutrosophic Graph Cut (NGC) is proposed for an efficient image segmentation algorithm.

In order to evaluate the performance of these proposed algorithms and methods, experimental studies on real datasets, artificial datasets, and on the various images have been done. Experimental studies show that the proposed methods have a robust and efficient structure, on images and datasets which contain uncertainty such as noise and outlier data points.

This thesis was supported by “The Scientific and Technological Research Center of Turkey” (TÜBİTAK), The Department of Science Fellowship and Grant Programmes (BİDEB), within the scope of “2214/A Doctorate Research Fellowship Program”.

Keywords: Neutrosophy, neutrosophic logic, neutrosophic set, image segmentation, pattern recognition, neutrosophic c-means, extreme learning machines, k-nearest neighbor.

ŞEKİLLER LİSTESİ

Sayfa No

TABLOLAR LİSTESİ

Sayfa No

KISALTMALAR LİSTESİ

2B : 2 boyutlu
AAÖM : Ağırlıklandırılmış aşırı öğrenme makinesi
AÇA : Ağırlıklandırılmış çevrimiçi ardışık
AGÇAT : Alt grup çevrimiçi ardışık topluluk
A*k*-EYK : Ağırlıklandırılmış *k*-en yakın komşuluk
ALÖÇT : Alt örnekleme tabanlı çevrimiçi torbalama
AÖM : Aşırı öğrenme makinesi
AŞÖÇT : Aşırı örnekleme tabanlı çevrimiçi torbalama
BCO : Bulanık c-ortalamalar
BÖ : Başarım ölçüsü
ÇA*k*-EYK : Çift ağırlıklandırılmış *k*-en yakın komşuluk
ÇBCO : Çekirdek bulanık c-ortalamalar
ÇÇK : Çekirdek çizge kesim
ÇK : Çizge kesim
ÇNCO : Çekirdek nötrozofik c-ortalamalar
DN : Doğru negatif
DP : Doğru pozitif
DVM : Destek vektör makinesi
EAA : Eğri altındaki alan
EKK : En küçük kesim
EKKA : En küçük kapsayan ağaç
EKY : En kısa yol
İKCO : İlişkisel kanıtlayıcı c-ortalamalar
***k*-EYK** : *k*-en yakın komşuluk
KCO : Kanıtlayıcı c-ortalamalar
KO : K-ortalamalar
MRA : Markov rassal alanlar
NAAÖM : Nötrozofik ağırlıklandırılmış aşırı öğrenme makinesi
NBK : Nötrozofik benzerlik kesim
NCO : Nötrozofik c-ortalamalar
NÇK : Nötrozofik çizge kesim

NK : Nötrozofik küme

NKCO : Nötrozofik kanıtlayıcı c-ortalamalar

NK-*k*-EYK : Nötrozofik küme *k*-en yakın komşuluk

Nkesim : Normalize kesim

OCO : Olasılı c-ortalamalar

RAÖ : Rastgele alt örnekleme

RTF : Radyal tabanlı fonksiyon

RY : Rastgele yürüyüş

SAAÖT : Sentetik azınlık alt örnekleme tekniği

SÇMK : Spektral çoklu manifold (katman) kümeleme

SGO : Sinyal gürültü oranı

SK : Spektral kümeleme

TGKİB : Tek-gizli katmanlı ileri beslemeli

YN : Yanlış negatif

YP : Yanlış pozitif

YSH : Yanlış sınıflandırma hatası

SİMGELER LİSTESİ

T	: Doğruluk üyelik değeri
I	: Belirsizlik üyelik değeri
F	: Yanlışlık üyelik değeri
N	: Veri (örnek) sayısı
C	: Küme merkezlerinin sayısı
x_i	: Veri noktası
$\bar{c}_{imax}$	: i veri noktası için hesaplanan küme merkezi
$\varpi_1, \varpi_2, \varpi_3$	: Ağırlıklandırma parametreleri
c_j	: j kümesinin merkezi
m	: Sabit bir sayı
δ	: Düzenleyici parametre
J	: Amaç (maliyet) fonksiyonu
ϕ	: Çekirdek fonksiyonu
K	: Çekirdek fonksiyonu
$L(.)$	: Lagrange fonksiyonu
λ	: Lagrange çarpanı
μ	: Ortalama
σ	: Standart sapma
$g(.)$	: Aktivasyon fonksiyonu
L	: Gizli düğüm sayısı
o_i	: Çıkış düğümü
t_i	: Çıkış vektörü
β	: Çıkış ağırlık vektörü
a_j	: Ağırlık vektörü
b_j	: Gizli düğümün eşik değeri
H	: Gizli katman çıkış matrisi
H^+	: H matrisinin Moore-Penrose genelleştirilmiş tersi
$\hat{\beta}$	: β'nın en küçük kareler çözümü
ξ	: Hata vektörü
W	: Ağırlıklar matrisi

ε : Sonlandırma kriteri

U : Üyelik değeri

d_i : Uzaklık fonksiyonu

1. GİRİŞ

1.1. Literatür İncelemesi

Bu tez çalışmasındaki literatür incelemesi, nötrozofik mantık çerçevesinde görüntü bölütleme ve örüntü tanıma alanları için bu çalışmada önerilen yöntemlerle ilgili konu başlıklarına göre yapılmıştır. Bu konu başlıkları: kümeleme, ağırlıklandırılmış aşırı öğrenme makinesi, *k*-en yakın komşuluk ve çizge kesim ile görüntü bölütlemedir.

1.1.1. Kümeleme

Veri kümelemesi veya küme analizi, örüntü tanıma ve makine öğrenmesinde önemli bir araştırma alanı olup, daha ileri uygulamalar için bir veri yapısının anlaşılmasına yardımcı olmaktadır. Kümeleme işlemi genellikle verilerin farklı kümelere bölünmesiyle ele alınmaktadır; burada kümeler içindeki benzerlik ve farklı kümeler arasındaki benzeşmezlik yüksektir.

K-Ortalamalar (KO) kümeleme, çok sayıda uygulamasıyla alanda öncü bir algoritma olarak bilinmektedir. Şimdiye kadar KO kümeleme algoritmasının birçok çeşidi önerilmiştir [1]. KO algoritması doğası gereği tüm veri noktalarına keskin üyelikler atamaktadır. Bulanık küme teorisi Zadeh tarafından [2] ortaya atıldıktan sonra keskin üyelikler kullanmak yerine üyelik fonksiyonları tarafından tanımlanan kısmi üyelikler kullanmak küme analizi açısından uygun sonuçlar verdiği görülmüştür.

Veri kümelemeye bulanık fikrini ilk önce Ruspini adapte etmiştir [3]. Dunn, yeni bir amaç fonksiyonunun yeniden tanımlandığı ve üyeliklerin mesafeye göre güncellendiği popüler Bulanık C-Ortalamalar (BCO) algoritmasını önermiştir [4]. Genelleştirilmiş BCO Bezdek tarafından tanıtılmıştır [5]. Birçok uygulamada başarılı sonuçlar almasına rağmen BCO'nun bazı sakıncaları vardır. Örneğin BCO, tüm veri noktalarının aynı öneme sahip olduğunu düşünmektedir. Gürültü ve aykırı veri noktaları da BCO'nun üstesinden gelemediği konulardır. Bu sakıncaları hafifletmek için geçmişte çeşitli girişimlerde bulunulmuştur. Gustafson ve Kessel çalışma [6]'da, farklı küme şekillerinin etkisini analiz etmek için BCO'da Mahalanobis mesafesini düşünmüşlerdir. Dave ve arkadaşları [7] dairesel ve eliptik biçimli veri kümeleri üzerinde etkili olan "Bulanık C-Kabuk" adlı yeni bir kümeleme algoritması önermiştir. Diğer taraftan, BCO'nun gürültüye karşı sakıncaları

incelenmiştir. Çalışma [8]'de araştırmacılar, BCO toplamının kısıtını bire gevşeterek ele alan Olasılı C-Ortalamalar (OCO) algoritması önerilmiştir. Pal ve arkadaşları [9] OCO ve BCO algoritmalarının bir kombinasyonu olarak düşünülen kümelenme merkezlerinin hem göreceli hem de mutlak benzerliklerini hesaba katmıştır.

Son zamanlarda, bir veri noktasının aynı anda birkaç alt kümeye ait olabileceğini düşünen sayısız kümeleme algoritmaları geliştirilmiştir [10]. Bu yaklaşımlar kanıtlayıcı (evidential) teorisine dayanarak benimsenmiştir [11]. Masson ve Denoeux, kanıtlayıcı c-ortalamalar algoritmasını (KCO) önermiştir [11]. Araştırmacılar, daha sonra ilişkisel KCO (İKCO) algoritmasını geliştirdiler [12]. Guo ve Şengür, nötrozofik mantığa dayalı olarak nötrozofik c-ortalamalar (NCO) [13] ve nötrozofik kanıtlayıcı c-ortalamalar (NKCO) [14] kümeleme algoritmalarını önermiştir. NCO'da, BCO yönteminin gürültü ve aykırı veriler üzerindeki zayıflığının üstesinden gelmek için yeni bir maliyet fonksiyonu geliştirilmiştir. NCO algoritmasında hem gürültü hem de aykırı değer reddi için iki yeni ret türü geliştirilmiştir.

BCO algoritmasının bir diğer önemli dezavantajı, doğrusal olmayan ayrılabilir kümelere karşı kümeleme başarısızlığıdır. Bu dezavantaj, BCO algoritmasındaki Mercer çekirdeği göz önüne alarak veri noktalarını doğrusal olmayan bir şekilde daha yüksek bir boyutsal öznitelik uzayına dönüştürerek azaltılabilir [15]. Mercer çekirdeğinin BCO algoritmasına eklenmesi çekirdek BCO (ÇBCO) algoritması olarak adlandırılmıştır ve özellikle dairesel ve eliptik şekilli kümeler üzerinde daha iyi kümeleme sonuçları elde etmiştir.

1.1.2. Ağırlıklandırılmış Aşırı Öğrenme Makinesi

Aşırı öğrenme makinesi (AÖM), 2016 yılında Huang ve arkadaşları [16] tarafından tek-gizli katmanlı ileri beslemeli (TGKİB) ağ olarak ileri sürülmüştür. AÖM'nin gizli katman parametreleri rastgele başlatılır ve çıkış ağırlıkları en küçük kareler algoritması kullanılarak belirlenir. Bu özellikten ötürü, AÖM hızlı öğrenme yeteneğine, daha iyi başarıma ve verimli hesaplama maliyetine sahiptir [16–19] ve sonuç olarak farklı alanlarda uygulanmıştır.

Bununla birlikte, AÖM büyük ve yüksek boyutlu gerçek veri kümesindeki alakasız değişkenlerin varlığından kötü şekilde etkilenmektedir [17, 20]. Metin sınıflandırması, hata tespiti, dolandırıcılık tespiti, uydu görüntülerinde petrol sızıntılarının tespiti, zehir bilimi, zirai modelleme ve tıbbi tanı gibi gerçek uygulamalarda dengesiz veri kümesi sorunu ortaya

çıkmaktadır [21]. Birçok zorlayıcı (zorlu) gerçek problemler, en az bir sınıfın diğerlerine göre az temsil edildiği eğitim verileriyle tanımlanmaktadır. Dengesiz veri sorunu genellikle, farklı sınıfların öğelerini yanlış sınıflandırmanın asimetrik maliyetiyle ilişkilendirilmiştir. Buna ek olarak, test veri kümesinin dağılımı eğitim örneklerinden farklı olabilir. Sınıf dengesizliği, bir sınıftaki örneklerin sayısı diğer sınıftan çok fazla olduğunda ortaya çıkmaktadır [22].

Dengesizlik sorunu ile mücadele etmeyi amaçlayan yöntemler, algoritmik tabanlı yöntemler, veri tabanlı yöntemler, maliyete duyarlı yöntemler ve sınıflandırıcı toplulukları temelli yöntemler olmak üzere dört gruba ayrılabilir [23]. Algoritmik tabanlı yaklaşımlarda, azınlık sınıfı sınıflandırma doğruluğu her sınıf için ağırlıkları ayarlayarak geliştirilmiştir [24]. Yeniden örnekleme yöntemleri, sınıflandırıcıların gelişimine katkı sağlamayan veri temelli yaklaşımlarda görülebilir [25]. Maliyet duyarlı yaklaşımlar, çoğunluk sınıfı ve azınlık sınıfının eğitim örneklerine çeşitli maliyet değerleri atamaktadır [26]. Son zamanlarda, topluluk temelli yöntemler dengesiz bir veri kümesinin sınıflandırılmasında yaygın olarak kullanılmaktadır [27]. Torbalama (bagging) ve artırma (boosting) yöntemleri, iki popüler topluluk yöntemidir.

Sınıf dengesizliği sorunu literatürde çok dikkat çekmiştir [28]. Sentetik azınlık aşırı örnekleme tekniği (SAAÖT) [24], azınlık sınıfı örneklerini yapay olarak elde etmek için ön işlemeyi kullanan en popüler yeniden örnekleme yöntemi olarak bilinmektedir. SAAÖT, her azınlık sınıfı örneği için en yakın azınlık sınıfı komşusuna katılan yeni bir örnek oluşturur. Sınırdaki-SAAÖT [29], Artırımlı-SAAÖT [30] ve Değiştirilmiş-SAAÖT [29], SAAÖT algoritmasının iyileştirilmiş (geliştirilmiş) biçimlerinden bazılarıdır. Buna ek olarak, sınıflandırılması zor bazı azınlık sınıfı örnekleri tanımlayan bir aşırı örnekleme yöntemi önerilmiştir [31]. Aşırı örneklemeyle torbalamanın kullanıldığı bir diğer aşırı örnekleme yöntemi [32]'de sunulmuştur. Araştırmacılar [33]'te, torbalama ve güçlendirme yöntemlerini birleştirerek çifte topluluk sınıflandırıcı kullanmayı seçmişlerdir. Başka bir çalışmada [34] araştırmacılar, çarpık (skewed) dağılıma sahip verilerin sınıflandırma başarımını artırmak için örnekleme ve topluluk tekniklerini birleştirmişlerdir. Eğitim seti dengeleninceye kadar çoğunluk sınıf örneklerini rastgele kaldıran bir başka yöntem, yani rastgele alt örnekleme (RAÖ) önerilmiştir [34]. Wang ve Japkowicz [35]'deki çalışmalarında saf Destek vektör makinesine (DVM) kıyasla azınlık sınıfı sınıflandırma doğruluğunun arttığı bir Artırımlı DVM önermişlerdir. Araştırmacılar tarafından [36]'da, k-en yakın komşuluk (*k*-EYK) sınıflandırıcısının kabul edildiği maliyete duyarlı bir yaklaşım

önerilmiştir. Buna ek olarak, dengesiz veri sınıflandırması için bir DVM tabanlı maliyete duyarlı bir yaklaşım [37]'de önerilmiştir. Dengesiz veri sınıflamasını ele almak için karar ağaçları [38] ve lojistik regresyon [39] tabanlı yöntemler de önerilmiştir.

Dengesiz bir veri kümesiyle eğitilmiş bir AÖM sınıflandırıcısı, çoğunluk sınıfına doğru önyargılı olabilir ve azınlık sınıfının doğruluğundan ödün vermek suretiyle çoğunluk sınıfında yüksek bir doğruluk elde edebilir. Ağırlıklandırılmış AÖM (AAÖM), dengesiz veri kümeleri üzerinde AÖM'nin sınıflandırma eksikliğini hafifletmek için kullanılmıştır ve maliyetle orantılı ağırlıklandırılmış örnekleme yöntemlerinden biri olarak görülebilir [40]. AÖM, iki-sınıflı bir problemdeki pozitif ve negatif gibi tüm veri noktalarına aynı yanlış sınıflandırma maliyet değerini atamaktadır. Negatif örneklerin sayısı pozitif örneklerin sayısından çok daha fazla olduğunda ya da tam tersi durumda, tüm örneklere aynı yanlış sınıflandırma maliyet değerinin atanması, geleneksel AÖM'nin dezavantajlarından biri olarak görülebilir. Doğrudan bir çözüm, sınıftaki dağılıma göre yanlış sınıflandırma maliyet değerlerini, sınıftaki örnek sayısıyla ters orantılı bir ağırlıklandırma şeması biçiminde elde etmektir. Araştırmacılar [22]'deki çalışmada yığın ile yığın (chunk-by-chunk) ve tek-tek (one-by-one) öğrenmede dengesizlik problemini hafifletmek için ağırlıklandırılmış çevrimiçi ardışık AÖM (AÇA-AÖM) algoritması önermişlerdir. Hesaplama açısından verimli bir şekilde ağırlık ayarı seçilmiştir. Tanimato ağırlıklandırılmış AÖM (T-AAÖM), kimyasal bileşik biyolojik etkinliğini ve diğer ayrık ikili gösterimi olan veriyi tahmin etmek için kullanılmıştır [41]. Araştırmacılar [42]'de el yazısı rakamları tanıması için bir TGKİB AÖM sunmuşlardır. Giriş ve çıkış ağırlıkları genel olarak en küçük karelerin toplu öğrenme türü ile optimize edilmiştir. Öznitelikler belirlenen pozisyonlara atanmıştır. Başka bir AAÖM algoritması olan alt-grup çevrimiçi ardışık topluluk AÖM (AGÇAT-AÖM), Mirza ve arkadaşları [43] tarafından AÇA-AÖM'den esinlenerek önerilmiştir. AGÇAT-AÖM, kavram sürükleyici bir veri akışından sınıf dengesizliği öğrenmeyi ele almayı amaçlamaktadır. Başka bir topluluk temelli AAÖM yöntemi Zhang ve arkadaşları [44] tarafından önerilmiştir; burada topluluktaki her temel öğrenicinin ağırlığı diferansiyel evrim algoritması ile optimize edilmiştir. Araştırmacılar [45], sınıf dengesizliğini öğrenmek için aşırı-örnekleme-tabanlı çevrimiçi torbalama (AŞÖÇT) ve alt-örnekleme-tabanlı çevrimiçi torbalama (ALÖÇT) kapsamında yeniden örnekleme stratejisini daha da geliştirmişlerdir.

Dengesizlik konusunda çok fazla farkındalık yaratılmış olmasına rağmen, önemli (kilit) konuların birçoğu hala çözülmemiştir ve büyük veri kümelerinde daha sık

karşılaşılmaktadır. Ağırlık değerlerinin belirlenmesi AÖM tasarımının anahtarıdır. Gürültü ve aykırı veriler gibi farklı durumlar dikkate alınmalıdır.

1.1.3. *k*-En Yakın Komşuluk

En eski ve en basit yaklaşım olarak bilinen *k*-en yakın komşuluk (*k*-EYK) parametrik olmayan eğiticili bir sınıflandırıcıdır [46, 47]. Bu sınıflandırıcı, bilinmeyen bir örneğin sınıf etiketini, eğitim setinde depolanan *k*-EYK ile belirlemeyi amaçlamaktadır. En yakın *k* komşulukları bazı uzaklık fonksiyonlarına dayalı olarak belirlenmektedir. En eski ve en basit yaklaşım olduğundan, kalp ritim bozukluğu bulma [48], iflas tahmini [49], diyabet hastalıklarının teşhisi [50], insan eyleminin tanınması [51] metin sınıflandırması [52] gibi çok sayıda veri madenciliği ve örüntü tanıma uygulaması yapılmıştır.

k-EYK başarılı sonuçlar vermesine rağmen, onun hassasiyetini artırmak için bazı iyileştirmeler yapılmıştır. Bulanık teori tabanlı *k*-EYK (Bulanık *k*-EYK) en başarılı olanlar arasındadır. *k*-EYK, eğitim veri örnekleri için net (keskin) üyelik değerleri üretirken bulanık *k*-EYK, sınıf etiketlerinin belirlenmesini artıran sürekli üyelik değerlerini net üyelik değerlerinin yerine kullanmaktadır.

Bulanık teoriyi *k*-EYK yaklaşımına dahil eden kişiler Keller ve arkadaşlarıdır [53]. Araştırmacılar, bulanık üyelikleri etiketlenmiş örneklere atamak için üç farklı yöntem önermişlerdir. Bulanık üyeliklerin belirlenmesinden sonra, test örneğinin nihai sınıf etiketinin belirlenmesi için bulanık üyeliklerin ağırlıklandırılması gerekmektedir bunun için de bazı uzaklık fonksiyonları kullanılmıştır. Geleneksel bulanık *k*-EYK algoritması tarafından üyelik atanması, bazı uzaklık fonksiyonlarının seçimine bağlı olduğu için bir dezavantaja sahiptir. Bu dezavantajı gidermek için, Pham ve arkadaşları [54] en uygun şekilde ağırlıklandırılmış bir bulanık *k*-EYK yaklaşımı önermişlerdir. Araştırmacılar, bulanık *k*-EYK yaklaşımının etkinliğini artırmak için kullanılan en uygun ağırlıkları belirlemek için bir hesaplama şeması tanıtmışlardır.

Denoeux ve arkadaşları [55], eğitim veri örneklerinin üyeliklerini hesaplamak için Dempster-Shafer teorisinin kullanıldığı bir *k*-EYK yöntemi önermişlerdir. Araştırmacılar, sınıflandırılacak bir örneğin her bir komşusunun bir kanıt maddesi olarak kabul edildiğini ve destek derecesinin, uzaklığın bir fonksiyonu olarak tanımlandığını varsaymışlardır. Nihai sınıf etiket ataması Dempster'in kombinasyon kuralıyla halledilmiştir. Diğer bir kanıt teorisine dayalı *k*-EYK yaklaşımı Zouhal ve arkadaşları tarafından önerilmiştir [56]. Araştırmacılar, aitlik derecesine ek olarak, belirsizliği modellemek için bilgisizlik sınıfını

geliştirmişlerdir. Sonrasında, Zouhal ve arkadaşları [57], BK*k*-EYK ile gösterilen genelleştirilmiş bir *k*-EYK yaklaşımı önermiştir. Böylece araştırmacılar *k*-EYK'nın sınıflandırma başarımını artırmak için bulanık teoriyi benimsemişlerdir. BK*k*-EYK fikri, her bir eğitim örneğinin her sınıf için bir dereceye kadar üyeliğe sahip olduğu düşüncesinden ortaya çıkmıştır. Buna ek olarak, Liu ve arkadaşları [58], kanıt bulmaya dayalı bulanık-inanç (fuzzy-belief) *k*-en yakın komşuluk (Bİ*k*-EYK) sınıflandırıcıyı önermişlerdir. Bİ*k*-EYK algoritmasında her bir etiketlenmiş örnek kendi komşuluklarına göre her sınıfa bir bulanık üyelikle atanmıştır. Test örneğinin sınıf etiketi, nesne ile k en yakın komşuları arasındaki uzaklıklardan belirlenen k temel inanç (belief) atamaları ile belirlenmiştir.

İ*k*-EYK sınıflandırıcısı ile gösterilen *k*-EYK tabanlı bir inanç teorisi Liu ve arkadaşları tarafından tanıtılmıştır [59]. Araştırmacılar meta sınıfı kullanarak belirsiz verileri ele almayı hedeflemişlerdir. Önerilen yöntem başarılı sonuçlar üretse de, hesaplama karmaşıklığı ve *k*'ye olan duyarlılık, birçok sınıflandırma uygulaması için yaklaşımı uygun bulmamaktadır.

Derrac ve arkadaşları [60], aralık değerli (interval-valued) bulanık kümelerin kullanıldığı bir evrimsel (evolutionary) bulanık *k*-EYK yaklaşımı önermiştir. Araştırmacılar, sadece yeni bir üyelik fonksiyonu önermekle kalmamışlar, aynı zamanda yeni bir oylama şeması önermişlerdir. Dudani ve arkadaşları [61], uzaklıkla ağırlıklandırılmış (distance-weighted) *k*-EYK (A*k*-EYK) olarak adlandırılan *k*-EYK için ağırlıklandırılmış bir oylama yöntemi önermiştir. Araştırmacılar, uzaklıkla ağırlıklandırılmış fonksiyonu kullanarak yakın komşuların uzak olanlardan daha fazla ağırlık aldığını varsaymışlardır. Gou ve arkadaşları [62], uzaklıkla çift ağırlıklandırılmış bir fonksiyonun ortaya konduğu bir uzaklıkla ağırlıklandırılmış *k*-EYK (ÇA*k*-EYK) yöntemini önermişlerdir. Önerilen yöntem, *k* değerinin seçimi için yeni bir yaklaşım kullanarak geleneksel *k*-EYK'nın başarımını geliştirmiştir.

1.1.4. Çizge Kesim ile Görüntü Bölütleme

Klasik bir tanımla görüntü bölütlemesi, bir giriş görüntüsünün önceden tanımlanmış kriterlere göre birbirinden ayrı, homojen ve anlamlı alt görüntülere bölünmesi işlemidir. Görüntü bölütleme ayrıca birçok bilgisayar görmesi ve örüntü tanıma uygulamalarında önemli ve kritik bir adım olarak bilinmektedir. Birçok araştırmacı görüntü bölütleme üzerine çalışmaktadır ve çok sayıda çalışma yapılmıştır [63].

Yayınlanmış çalışmalar arasında, çizge tabanlı bölütleme algoritmaları önemli bir görüntü bölütleme kategorisini oluşturmaktadır [64]. Bir G çizgesi, $G = (V, E)$ olarak

gösterilebilir; burada V köşelerin bir kümesi ve E kenarların bir kümesidir. Bir görüntüde köşeler bir piksel veya bir bölge olabilir ve kenarlar komşu köşeleri birbirine bağlamaktadır [65]. Ağırlık, piksellerin bazı özelliklerini kullanarak her bir kenar ile ilişkilendirilen negatif olmayan bir benzeşmezlik ölçütüdür.

Nötrozofinin görüntü üzerindeki belirsizliği yorumlama avantajı kullanılarak görüntü bölütleme için NK ve çizge kesim (ÇK) birleştirilmiştir. NK bulanık kümenin bir uzantısıdır [66]. NK teorisinde bir kümenin üyelerinin sırasıyla doğruluk, yanlışlık ve benzerlik dereceleri vardır [67]. Bu nedenle, belirsizlik bilgileriyle baş edebilme yeteneğine sahiptir ve hemen hemen tüm mühendislik topluluklarında dikkat çekmiştir. Daha sonra, NK tabanlı renk ve doku bölütlemesi [68–75], NK tabanlı kümeleme [13, 76, 77], görüntü eşiklemesi için NK tabanlı benzerlik [78], NK tabanlı kenar algılama [79] ve NK tabanlı seviye ayarı (level set) [80] gibi çok sayıda çalışma yapılmıştır.

Peng ve arkadaşları [81] tarafından çizge tabanlı görüntü bölütleme üzerine sistematik bir anket çalışması yapılmıştır. Bu çalışmada, araştırmacılar çizge tabanlı görüntü bölütleme yöntemlerini beş gruba ayırmıştır. İlk kategori en küçük kapsayan ağaç (EKKA) yöntemidir. EKKA, çizge teorisinde sayısız eser ile popüler bir kavramdır. Çalışma [82]'de EKKA'ya dayalı hiyerarşik bir görüntü bölütleme yöntemi önerilmiştir. Bu yöntem giriş görüntüsünü yinelemeli bir şekilde bölütlemiştir. Her yinelemelede, bir alt çizge oluşturulmuş ve son yinelemede belirli sayıda alt çizge vardır. Çalışma [83]'te, iki alt çizgedeki ve çizgeler arasındaki farkları kullanarak EKKA tabanlı bir görüntü bölütleme algoritması üretmek için bir bölge birleştirme işlemi benimsenmiştir.

Maliyet fonksiyonu tabanlı çizge kesim yöntemleri ikinci kategoriyi oluşturmaktadır. En popüler çizge tabanlı bölütleme yöntemleri bu kategoridedir. Wu ve arkadaşları [65] çizge teorisini görüntü bölütlemeye uygulamıştır ve bir maliyet fonksiyonunu en aza indirgemek için popüler bir en küçük kesim (EKK) yöntemini önermiştir. Normalize kesim (Nkesim) adında bir çizge tabanlı görüntü bölütleme yaklaşımı sunulmuştur [84]. Bu yaklaşım, öz değerli sistem sunarak EKK yönteminin eksikliklerini hafifletmiştir. Wang ve arkadaşları [85] çizge tabanlı bir maliyet fonksiyonu ve yöntem sunmuştur ve bunu kesim sınırı boyunca farklı kenar ağırlıklarının toplamının oranı olarak tanımlamışlardır. Ding ve arkadaşları [86] iki alt çizge arasındaki benzerliğin en aza indirgendiği ve her bir alt çizge içindeki benzerliğin en yükseğe çıkarıldığı EKK yönteminin zayıflığını azaltmak için bir maliyet fonksiyonu sunmuştur. Başka bir etkin çizge tabanlı görüntü bölütleme yöntemi [87]'de önerilmiştir ve hem iç hem de sınır bilgileri dikkate alınmıştır. Dış sınır ve iç bölge

arasındaki oranı en aza indirgemiştir. Ortalama kesim, kesim sınırındaki ortalama kenar ağırlığını en aza indirgemek için kenar ağrılık fonksiyonunu [85] içermektedir.

Markov rassal alanlarına (MRA) dayanan yöntemler üçüncü kategoride ve en kısa yol (EKY) tabanlı yöntemler dördüncü kategoride bulunmaktadır. Genellikle MRA tabanlı çizge kesim yöntemleri maliyet fonksiyonlu bir çizge yapısını oluştururlar ve bölütleme problemini çözmek için bu maliyet fonksiyonunu en aza indirgemeye çalışırlar. EKY tabanlı yöntemler, iki köşe arasındaki en kısa yolu aramaktadır [81] ve bölümlerin sınırları en kısa yol kullanılarak elde edilmektedir. EKY tabanlı bölütleme yöntemleri kullanıcılarla etkileşime ihtiyaç duymaktadır.

Diğer çizge tabanlı yöntemler beşinci kategoride toplanmıştır. Grady'nin [88] rastgele yürüyüş (RY) yöntemi, piksel etiketlerini elde etmek için ağırlıklandırılmış bir çizge kullanmıştır ve sonra bu ağırlıklar RY'nin kenardan geçme ihtimali olarak değerlendirilmiştir. Son olarak bir piksel etiketi, RY'nin ilk ulaştığı tohum noktası tarafından atanmıştır.

1.2. Tezin Amacı

Bu tez çalışmasının amacı, nötrozofik mantık çerçevesinde görüntü bölütleme ve örüntü tanıma alanları için yeni yöntemler ve algoritmalar geliştirmektir. Nötrozofi, doğası gereği, belirsizlik durumlarının üstesinden gelebilmektedir. Bu amaçla belirsizlik durumunu giderebilen nötrozofik tabanlı yeni bir kümeleme algoritması önerilmiştir. Bunun yanı sıra nötrozofik kümenin katkısıyla belirsiz ve aykırı veriye sahip veri kümelerinde (sınıflarında) başarılı olması ön görülen iki adet yeni sınıflandırıcı yöntemi önerilmesi amaçlanmıştır. Önerilen bu sınıflandırıcıların literatürdeki birçok sınıflandırma probleminin çözümünde kullanılacağı ümit edilmektedir. Ayrıca görüntü bölütleme uygulaması olarak özellikle gürültü içeren görüntüler için yeni bir nötrozofik çizge kesim yöntemi geliştirilmesi düşünülmüştür.

1.3. Tezin Literatüre Katkıları

Bu tez çalışması kapsamında bir adet yurtiçi uluslararası konferans bildirisi [89] sunulmuş ve beş adet uluslararası *Science Citation Index Expanded* indeksli makale [90–94] yayınlanmıştır. Ayrıca bu tez çalışması, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) Bilim İnsanı Destekleme Başkanlığı (BİDEB) tarafından verilen 2214/A Yurt

dışı doktora sırası araştırma burs programı kapsamında 2016/2 döneminde 9 ay süreyle desteklenmiştir.

Yayınlar ve araştırma bursuyla ilgili ayrıntılar aşağıda belirtilmiştir:

- Akbulut, Y., Şengür, A., & Guo, Y. (2016, September). Texture segmentation based on Gabor filters and neutrosophic graph cut. In *International conference on advanced technology & sciences (ICAT'16), Konya, Turkey* (pp. 336-339).
- Akbulut, Y., Şengür, A., Guo, Y., & Polat, K. (2017). KNCM: Kernel Neutrosophic c-Means Clustering. *Applied Soft Computing*, *52*, 714-724.
- Akbulut, Y., Şengür, A., Guo, Y., & Smarandache, F. (2017). A Novel Neutrosophic Weighted Extreme Learning Machine for Imbalanced Data Set. *Symmetry*, *9*(8), 142.
- Akbulut, Y., Şengür, A., Guo, Y., & Smarandache, F. (2017). NS-k-NN: Neutrosophic Set-Based k-Nearest Neighbors Classifier. *Symmetry*, *9*(9), 179.
- Guo, Y., Akbulut, Y., Şengür, A., Xia, R., & Smarandache, F. (2017). An Efficient Image Segmentation Algorithm Using Neutrosophic Graph Cut. *Symmetry*, *9*(9), 185.
- Guo, Y., Şengür, A., Akbulut, Y., & Shipley, A. (2018). An Effective Color Image Segmentation Algorithm Based on Neutrosophic Adaptive Meanshift Clustering. *Measurement*, *119*, 28-40.
- TÜBİTAK BİDEB 2214/A yurt dışı doktora sırası araştırma bursu 2016/2 döneminde Department of Computer Science University of Illinois at Springfield (UIS), Springfield, IL, USA bölümü için destek alınması sonrası Eylül 2017 ve Haziran 2018 tarihleri arasında araştırma çalışması yapılmıştır.

1.4. Tezin Organizasyonu

Bu doktora tezi yedi bölümden oluşmaktadır. Birinci bölümde, tez konusuyla ilgili kapsamlı bir literatür araştırması ve incelemesi yapılmıştır. Tezin amaçları ortaya konarak literatüre yapılan katkılar belirtilmiştir. Son olarak tezin organizasyonu yine bu bölümde verilmiştir.

İkinci bölümde, tez çalışması süresince tez konusu ile ilgili olarak ele alınan ve çalışılan yöntemlerin, algoritmaların temelini oluşturan *Nötrozofi, Nötrozofik Küme, Nötrozofik Mantık, Nötrozofik Görüntü Bölütleme* kavramları ve konuları hakkında bilgi verilmiştir.

Üçüncü bölümde, veri kümelemedeki zorlukların başında gelen gürültü ve aykırı veri ele alınmıştır. Bu amaçla literatürde mevcut olan ve bu tür belirsizlik durumlarına karşı dayanıklı olan nötrozofik c-ortalamalar (NCO) algoritmasından esinlenerek çekirdek nötrozofik c-ortalamalar (ÇNCO) yöntemi geliştirilmiştir. Böylece hem belirsizlik durumuyla hem de doğrusal olmayan ayrılabilir veriler üzerinde kümeleme çalışmaları yapılmıştır. ÇNCO yöntemi, yapay ve gerçek veri kümelerinde test edilmiştir. Bu bölümdeki çalışmaların değerlendirilmesi sonucunda bir makale yayınlanmıştır [90].

Dördüncü bölümde, dengesiz veri kümelerindeki sınıflandırmada yetersiz kalan aşırı öğrenme makinesinin (AÖM) nötrozofik açıdan yorumlanmasıyla yeni bir nötrozofik ağırlıklandırılmış AÖM (NAAÖM) yöntemi ortaya çıkmıştır. NAAÖM yönteminin değerlendirilmesi amacıyla yapay veri kümeleri ve gerçek veri kümeleri üzerinde sınıflandırma uygulamaları yapılmıştır. Bu bölümdeki çalışmaların değerlendirilmesi sonucunda bir makale yayınlanmıştır [91].

Beşinci bölümde, Nötrozofik Küme tabanlı *k*-En Yakın Komşuluk (NK-*k*-EYK) sınıflandırıcısı tasarlanmıştır ve önerilmiştir. Sınıflandırıcının, sınıflandırma başarım oranını artırmak için NK üyelikleri benimsenmiştir. NK-*k*-EYK yönteminin başarımı yapay ve gerçek dünya veri kümeleri üzerinde kapsamlı deneyler yapılarak değerlendirilmiştir. Bu bölümdeki çalışmaların değerlendirilmesi sonucunda bir makale yayınlanmıştır [92].

Altıncı bölümde, nötrozofinin görüntü üzerindeki belirsizliği yorumlama avantajı kullanılarak görüntü bölütleme için NK ve çizge kesim (ÇK) birleştirilmiş ve etkili görüntü bölütlemesi için Nötrozofik Çizge Kesim (NÇK) adında yeni bir yöntem önerilmiştir. Bu yöntemde, uzamsal ve yoğunluk bilgisine dayalı bir belirsizlik filtresi oluşturulmuş ve görüntü üzerinde bir çizge tanımlanmıştır. Görüntüleri bölütlemek için en büyük akış algoritmasından yararlanılmıştır. Hem gürültüsüz hem de farklı gürültü seviyelerine sahip doğal görüntüler kullanılarak NÇK yöntemi değerlendirilmiştir. Bu bölümdeki çalışmaların değerlendirilmesi sonucunda bir makale [93] ve bir bildiri [89] yayınlanmıştır.

Son bölüm olan yedinci bölümde, tez çalışması ve önerilen yöntem ve algoritmaların sonuçları değerlendirilmiş ve katkıları belirtilmiştir. Gelecekte çalışılması düşünülen konulara yer verilmiştir.

2. NÖTROZOFİ VE NÖTROZOFİK MANTIK

2.1. Giriş

Nötrozofi, felsefenin, mantığın, küme teorisinin, olasılık ve istatistik bilgisinin bir araya getirildiği bir felsefe dalıdır [67]. Nötrozofi, belirsizliği temsil eden <Neut-A> adlı yeni bir kavram sunmaktadır ve bu bulanık mantıkla çözülemeyen bazı problemleri çözebilmektedir [71]. Örnek olarak, bir makale iki hakeme gönderilmiş olsun ve her ikisi de makalenin kabulünü %90 olarak belirlesinler. Ancak iki hakemin farklı geçmişleri olabilir. Bunlardan birisi bir uzman ve diğeri de bu alandaki yeni bir kişi olsun. Aynı kabul düzeyine sahip olmalarına rağmen, iki hakemin makalenin son kararı üzerindeki etkileri farklı olmalıdır. Bulanık mantığın iyi bir iş çıkaramadığı belirsiz koşulları içeren, hava durumu tahmini, hisse senedi fiyatı tahmini ve siyasi seçimler gibi birçok benzer problem bulunmaktadır [2].

2.2. Nötrozofi

Latince'den tarafsız anlamında "neuter" ve Yunanca'dan beceri/bilgelik anlamında "sophia" kelimelerinden oluşan "Neutrosophy" Nötrozofi, 1980 yılında Smarandache tarafından tanıtılmıştır [66]. Nötrozofi, t'nin doğru, i'nin belirsiz ve f'nin yanlış olduğu önermesine dayanan bulanık mantığın bir genelleştirilmesidir. t, i ve f; T, I ve F aralıklarından gerçek değerlerdir ve bunlar üzerinde hiçbir kısıtlama yoktur. Aşağıda farklı mantık türlerine örnek verilmektedir [66]:

- Sezgisel mantık: $0 < n < 100$ için, $0 \leq t, i, f \leq 100$.
- Bulanık mantık: $n = 100$ için, $i = 0$ ve $0 \leq t, f \leq 100$.
- İkili (Boolean) mantık: $n = 100$ için, $i = 0$, t ile f ya 0 ya da 100.
- Paraconsistent mantık: $n > 100$ için, $t, f < 100$.
- Dialetheist mantık: $t = f = 100$ ve $i = 0$.

Aşağıdaki iki örnek, bulanık mantığın genellemesi olan nötrozofinin, diğer mantık türlerine göre insan mantığına nasıl daha yakın olduğunu göstermeye yardımcı olacaktır:

- "Yarın yağmur yağacak" ifadesinde, sabit bir değerden bahsetmiyoruz. Nötrozofik terimlerle %60 doğru (t), %50 belirsiz (i) ve %30 yanlış (f) olduğunu söyleyebiliriz [95].
- Gerçeklik değeri de gözlemciye göre değişmektedir. "Yaman zekidir" ifadesi, danışmanına göre (0,35, 0,67, 0,6), kendisine göre (0,8, 0,25, 0,1), eşine göre (0,5, 0,2, 0,3) olabilir.

Nötrozofi insan aklına daha yakındır, çünkü insan zihni gibi, çeşitli gözlemcilerden aldığı bilginin ya da dilsel yanlışlığın belirsizliğini yakalar. Belirsizlik, eksik bilgiden, edinim hatalarından veya rastgelelikten kaynaklanabilir [96]. Nötrozofi, bulanık mantık, klasik mantık ve kesin olmayan olasılığı genelleştiren çoklu değerli bir mantık olan nötrozofik mantığın temelidir.

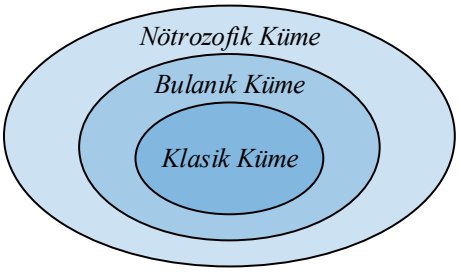
2.3. Nötrozofik Küme ve Nötrozofik Mantık

Bir Nötrozofik Küme (NK), sezgisel kümenin, bulanık kümenin, paraconsistent kümenin, dialetheist kümenin, paradoksist kümenin ve bir totolojik kümenin genelleştirilmesidir [2, 67, 97–99]. <A> bir varlık ya da olay olarak tanımlansın; bu durumda <Non-A>, <A> olmayanlar ve <Anti-A>, <A>'nın tersidir. Ayrıca <Neut-A>, ne <A> ne de <Anti-A> olarak tanımlanamaz. Örnek olarak, eğer <A> = Beyaz ise <Anti-A> = Siyah'tır. <Non-A> = Yeşil, Mavi, Sarı, Kırmızı, Siyah, vb. (Beyaz dışındaki herhangi bir renk). <Neut-A> = Yeşil, Mavi, Sarı, Kırmızı, vb. (Siyah ve Beyaz dışındaki herhangi bir renk) [95].

<A>, <Neut-A> ve <Anti-A>'yı temsil etmek için T, I, ve F nötrozofik bileşenler olarak tanımlansın. T, I, ve F; $sup\ T = t_{sup}$, $inf\ T = t_{inf}$, $sup\ I = i_{sup}$, $inf\ I = i_{inf}$, $sup\ F = f_{sup}$, $inf\ F = f_{inf}$, $n_{sup} = t_{sup} + i_{sup} + f_{sup}$ ve $n_{inf} = t_{inf} + i_{inf} + f_{inf}$ olmak üzere $]0^-, 1^+[$'in standart veya standart dışı gerçek alt kümeleridir [100]. x_{sup} alt kümelerin üst sınırlarını, x_{inf} ise alt kümelerin alt sınırlarını belirtmektedir. T, I, ve F mutlaka aralıklarla değil, ancak herhangi bir gerçek alt üniter bir alt küme olabilir. T, I, ve F bilinen veya bilinmeyen parametrelere bağlı olarak değerleri belirlenmiş vektör fonksiyonları veya işlemlerdir ve bunlar sürekli veya ayrık olabilirler.

Bir nötrozofik küme ile bulanık küme arasındaki en temel fark, bir bulanık kümede $n = t + f$ ifadesi 1'e eşit olmak zorunda iken bir nötrozofik kümedeki toplam n üzerinde herhangi bir sınırlama bulunmamasıdır [101].

Şekil 2.1'de NK ile bulanık küme ve klasik küme arasındaki ilişki gösterilmiştir.



Şekil 2.1. Nötrozofik küme, bulanık küme ve klasik küme arasındaki ilişki [95]

Burada kaynak [66]'daki çalışmada nötrozofik küme kavramlarına ait birkaç tanımlamaya yer verilmiştir.

Tanım 1 (Nötrozofik Küme): X noktaların (nesnelerin) bir uzayı olsun, X'deki genel bir öğe x ile gösterilir. X'deki bir A nötrozofik kümesi, bir gerçek üyelik fonksiyonu T_A, bir belirsiz üyelik fonksiyonu I_A ve bir yanlış üyelik fonksiyonu F_A ile tanımlanır. $T_A(x)$, $I_A(x)$ ve $F_A(x)$, $]0^-, 1^+[$'in gerçek standart veya standart dışı alt kümeleridir:

$$T_A: X \rightarrow]0^-, 1^+[\tag{2.1}$$

$$I_A: X \rightarrow]0^-, 1^+[\tag{2.2}$$

$$F_A: X \rightarrow]0^-, 1^+[\tag{2.3}$$

$T_A(x)$, $I_A(x)$ ve $F_A(x)$ 'in toplamları üzerinde bir kısıtlama yoktur, böylece $0^- \leq sup T_A(x) + \sup I_A(x) + sup F_A(x) \leq 3^+$.

Tanım 2 (Eşlenik): Bir A nötrozofik kümesinin eşleniği (tamamlayıcısı) $\bar{A}$ ile gösterilir ve X'deki bütün x'ler için (2.4), (2.5) ve (2.6)'daki gibi tanımlanır:

$$T_{\bar{A}}(x) = \{1^+\} \ominus T_A(x) \tag{2.4}$$

$$I_{\bar{A}}(x) = \{1^+\} \ominus I_A(x) \tag{2.5}$$

$$F_{\bar{A}}(x) = \{1^+\} \ominus F_A(x) \tag{2.6}$$

Tanım 3 (Birleşme): İki nötrozofik küme olan A ve B'nin birleşimi nötrozofik bir C kümesidir, $C = A \cup B$ olarak yazılır. Gerçek üyelik, belirsiz üyelik ve yanlış üyelik fonksiyonları (2.7), (2.8) ve (2.9)'da görüldüğü gibi A ve B'dekilerle ilgilidir.

$$T_C(x) = T_A(x) \oplus T_B(x) \ominus T_A(x) \odot T_B(x) \quad \textbf{(2.7)}$$

$$I_C(x) = I_A(x) \oplus I_B(x) \ominus I_A(x) \odot I_B(x) \quad \textbf{(2.8)}$$

$$F_C(x) = F_A(x) \oplus F_B(x) \ominus F_A(x) \odot F_B(x) \quad \textbf{(2.9)}$$

Tanım 4 (Kesişim): İki nötrozofik küme olan A ve B'nin kesişimi nötrozofik bir C kümesidir, $C = A \cap B$ olarak yazılır. Gerçek üyelik, belirsiz üyelik ve yanlış üyelik fonksiyonları (2.10), (2.11) ve (2.12)'de görüldüğü gibi A ve B'dekilerle ilgilidir.

$$T_C(x) = T_A(x) \odot T_B(x) \quad \textbf{(2.10)}$$

$$I_C(x) = I_A(x) \odot I_B(x) \quad \textbf{(2.11)}$$

$$F_C(x) = F_A(x) \odot F_B(x) \quad \textbf{(2.12)}$$

Tanım 5 (Fark): İki nötrozofik küme olan A ve B'nin farkı nötrozofik bir C kümesidir, $C = A \backslash B$ olarak yazılır. Gerçek üyelik, belirsiz üyelik ve yanlış üyelik fonksiyonları (2.13), (2.14) ve (2.15)'te görüldüğü gibi A ve B'dekilerle ilgilidir.

$$T_C(x) = T_A(x) \ominus T_A(x) \odot T_B(x) \quad \textbf{(2.13)}$$

$$I_C(x) = I_A(x) \ominus I_A(x) \odot I_B(x) \quad \textbf{(2.14)}$$

$$F_C(x) = F_A(x) \ominus F_A(x) \odot F_B(x) \quad \textbf{(2.15)}$$

2.4. Nötrozofik Kümelerle İşlemler

S_1 ve S_2 iki nötrozofik küme olması durumunda aşağıdaki işlemler tanımlanabilir [102]:

- Toplama

  $S_1 \oplus S_2 = \{x | x = s_1 + s_2, burada\ s_1 \in S_1\ ve\ s_2 \in S_2\}$

  $\{1^+\} \oplus S_2 = \{x | x = 1^+ + s_2, s_2 \in S_2\}$

  $inf\ S_1 \oplus S_2 = inf\ S_1 + inf\ S_2$

  $sup\ S_1 \oplus S_2 = sup\ S_1 + sup\ S_2$

- Çıkarma

  $S_1 \ominus S_2 = \{x | x = s_1 - s_2, burada\ s_1 \in S_1\ ve\ s_2 \in S_2\}$

  $\{1^+\} \ominus S_2 = \{x | x = 1^+ - s_2, s_2 \in S_2\}$

  $inf\ S_1 \ominus S_2 = \inf S_1 - \inf S_2$

  $sup\ S_1 \ominus S_2 = sup\ S_1 - sup\ S_2$

- Çarpma

  $S_1 \odot S_2 = \{x | x = s_1 \cdot s_2, burada\ s_1 \in S_1\ ve\ s_2 \in S_2\}$

  $inf\ S_1 \odot S_2 = inf\ S_1 \cdot inf\ S_2$

  $sup\ S_1 \odot S_2 = \sup S_1 \cdot \sup S_2$

- Kümenin bir sayıya bölünmesi

  $k \in R\ olsun. S_1 \oslash k = \{x | x = s_1/k, burada\ s_1 \in S_1\}$

2.5. Nötrozofik Piksel

Nötrozofinin uygulanabilmesi için görüntünün nötrozofik bir etki alanına aktarılması gerekmektedir. Nötrozofik alandaki bir piksel $P\{T, I, F\}$ olarak gösterilebilir; pikselin anlamı % t doğru, % i belirsiz ve % f yanlıştır. Burada sırasıyla t, T içinde; i, I içinde ve f, F içinde değişmektedir. Klasik bir kümede $i = 0$, t ve f ya 0 ya da 100'dür. Bulanık bir kümede $i = 0$ ve $0 \leq t, f \leq 100$'dür. Nötrozofik bir kümede $0 \leq t, i, f \leq 100$'dür.

2.6. Nötrozofik Görüntü Bölütleme

Nötrozofik görüntü bölütleme örneği olarak bir ultrason görüntüsündeki lezyonların arka plandan bölütlenmesi ele alınmıştır. Burada, ön plan (lezyonlar) ve arka plan olmak üzere iki sınıf bulunmaktadır [103].

Bir görüntüde, (u, v) konumundaki bir p pikseli için üç tane nötrozofik bileşen T, I ve F vardır ve şu şekilde tanımlanır:

$$T(u,v) = 1 - G(u,v) \tag{2.16}$$

$$I(u,v) = BULANIKLIK(u,v) * (1 - KENAR(u,v)) \tag{2.17}$$

$$F(u,v) = 1 - T(u,v) \tag{2.18}$$

burada

$$BULANIKLIK(u,v) = \begin{cases} 2\big(1 - T(u,v)\big) & T(u,v) \geq 0{,}5 \\ 2T(u,v) & T(u,v) < 0{,}5 \end{cases} \tag{2.19}$$

$$KENAR(u,v) = \begin{cases} 1 \;\; eğer\; p\; kenar\; ise \\ 0 \;\; eğer\; p\; kenar\; değilse \end{cases} \tag{2.20}$$

$G(u, v)$, $[0, 1]$ aralığında normalize edilmiş yoğunluk değeridir. Nötrozofik bileşen olarak üç bileşenden bahsedilmişti. Bunlardan ilki, piksel yoğunluk değerinin [0, 1] aralığında normalize edilmesi ve bu sonucun 1 den çıkarılmasıyla elde edilen T değeridir. Bunun nedeni görüntü içinde lezyonun karanlık olması ve arka planın parlak olmasıdır. F değeri, T değerinin tamamlayıcısıdır. T ve F kolayca ve basit bir şekilde tanımlanmasına rağmen, belirsiz I kümesini tanımlamak oldukça ilginç ve zorlayıcıdır. Buradaki hedef, pikselleri ön plan ve arka plan olarak iki sınıfa ayırmaktır. Ön plan, lezyonların bölgesini temsil etmektedir ve arka plan da lezyon dışında geriye kalan tüm pikselleri temsil etmektedir. Ön işleme tabi tutulduktan sonra görüntü, karanlık bir lezyon bölgesi, parlak bir arka plan ve arka plan üzerinde koyu gürültü bölgelerinden oluşmaktadır. Ayrıca yoğunluk değerleri ön plan ve arka plan arasında bir yerde bulunan bazı bölgelerde bulunmaktadır. Orta yoğunluktaki bu piksellerin nasıl bölütleneceği, bölütleme doğruluğunu büyük ölçüde

etkileyecektir. Bu tür piksellere yüksek belirsizlik atamak için (2.19) kullanılarak bulanıklık matrisi hesaplanır. 0,5 civarında yoğunluk değerine sahip piksellerin belirsizlik değeri yüksektir ve 0 veya 1'e yakın yoğunluk değerlerine sahip piksellerin belirsizlik değeri düşüktür. Bununla birlikte, orta yoğunluklu tüm piksellerin belirsizlik değeri yüksek olmamalıdır. Eğer orta yoğunluğa sahip bir piksel lezyonun kenarında bulunuyorsa, piksel belirsiz bir değere sahip olmamalıdır. Aksi takdirde kenar, yüksek belirsizliğe sahip pikselleri değiştirmek için kullanılan komşuluk ortalamasından dolayı bulanık olacaktır.

3. ÇEKİRDEK NÖTROZOFİK C-ORTALAMALAR KÜMELEMESİ

3.1. Giriş

Bu bölümde, Çekirdek Bulanık C-Ortalamalar (ÇBCO) algoritmasından esinlenerek doğrusal olarak ayrılamayan veri kümeleri üzerinde Nötrozofik C-Ortalamalar (NCO) yönteminin iyileştirilmesi ile yeni bir Çekirdek Nötrozofik C-Ortalamalar (ÇNCO) yöntemi önerilmiştir. Bunu yapabilmek için NCO içine Mercer çekirdeği eklenerek NCO yeniden formüle edilmiştir. Bu nedenle, gürültüye ve aykırılıklara karşı dayanıklı parametre tahmini yapmak için yeni bir amaç (maliyet) fonksiyonu üretilmiştir. Buna ek olarak, önerilen amaç fonksiyonunun minimize edilmesinden yeni üyelik ve prototip güncelleme denklemleri türetilmiştir. ÇNCO'nun üç üyelik fonksiyonu vardır: T, I ve F. Her bir veri noktası için sırasıyla T belirli kümeler için üyelik derecesini, I belirsiz kümeler için üyelik derecesini ve F aykırı kümeler için üyelik derecesini belirtmektedir. T, I ve F üyelik değerleri, kümeleme sürecinde tekrar tekrar hesaplandığından gürültüye ve aykırılıklara karşı dayanıklıdır.

Önerilen ÇNCO yöntemi, yapay veri kümesi kümeleme, gerçek veri kümesi kümeleme ve gürültülü görüntü bölütleme gibi çeşitli uygulamalara uygulanmıştır. Elde edilen sonuçlar ÇBCO yöntemi ile karşılaştırılmış ve önerilen ÇNCO yönteminin ÇBCO yönteminden daha iyi sonuç verdiği görülmüştür.

3.2. Nötrozofik C-Ortalamalar Kümeleme

Veri kümeleme, veri madenciliği ve makine öğrenmesinde önemli bir eğiticisiz yöntemdir. Giriş verilerini bazı benzerlik ölçütlerine bağlı olarak kategorilere sınıflandırır. K-Ortalamalar (KO) ve Bulanık C-Ortalamalar (BCO) gibi geleneksel kümeleme algoritmalarında, veri örnekleri, gürültüyü ve aykırı veri örneklerini dikkate almaksızın aynı öneme sahiptirler. Bununla birlikte bazı uygulamalarda, veri kümeleri içinde saptanması gereken gürültüler ve aykırılıklar olabilir. Son zamanlarda hem gürültüyü hem de aykırılıkları gidermek için Nötrozofik C-Ortalamalar (NCO) adı verilen yeni bir kümeleme algoritması önerilmiştir [13]. NCO, her iki derecenin belirlenmesini ve belirsiz kümeleri göz önüne almaktadır. Önerilen yeni amaç fonksiyonu ve üyelikler (3.1)'de verilmiştir:

$$J_{NCO}(T,I,F,c) = \sum_{i=1}^{N}\sum_{j=1}^{C}(\varpi_1 T_{ij})^m \left||x_i - c_j|\right|^2 + \sum_{i=1}^{N}(\varpi_2 I_i)^m \left||x_i - \bar{c}_{imax}|\right|^2 + \delta^2 \sum_{i=1}^{N}(\varpi_3 F_i)^m \quad \textbf{(3.1)}$$

burada T_{ij}, I_i ve F_i belirlenen kümelere, sınır bölgelerine ve gürültülü veri kümesine ait üyelik değerleridir. $0 < T_{ij}, I_i, F_i < 1$, eşitsizliği aşağıdaki denklemle karşılanmaktadır.

$$\sum_{j=1}^{c} T_{ij} + F_i + I_i = 1 \quad \textbf{(3.2)}$$

Her bir i veri noktası için T_{ij}'nin en büyük ilk iki değerine ait kümelerin merkezleri kullanarak $\bar{c}_{imax}$ hesaplanmaktadır.

$$\bar{c}_{imax} = \frac{c_{pi} + c_{qi}}{2} \quad \textbf{(3.3)}$$

$$p_i = \arg\max_{j=1,2,\ldots,c}(T_{ij}) \quad \textbf{(3.4)}$$

$$q_i = \arg\max_{j\neq p_i \cap j=1,2,\ldots,c}(T_{ij}) \quad \textbf{(3.5)}$$

burada m bir sabittir. p_i ve q_i, T'nin en büyük ilk iki değerine ait olan küme sayılarıdır. p_i ve q_i tanımlandığında $\bar{c}_{imax}$ hesaplanır ve değeri her bir i veri noktası için sabit bir sayıdır. İlgili üyelik fonksiyonları aşağıdaki gibi hesaplanır.

$$T_{ij} = \frac{\varpi_2\varpi_3(x_i - c_j)^{-\left(\frac{2}{m-1}\right)}}{\sum_{j=1}^{C}(x_i - c_j)^{-\left(\frac{2}{m-1}\right)} + (x_i - \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)} + \delta^{-\left(\frac{2}{m-1}\right)}} \quad \textbf{(3.6)}$$

$$I_i = \frac{\varpi_1\varpi_3(x_i - \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)}}{\sum_{j=1}^{C}(x_i - c_j)^{-\left(\frac{2}{m-1}\right)} + (x_i - \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)} + \delta^{-\left(\frac{2}{m-1}\right)}} \quad \textbf{(3.7)}$$

$$F_i = \frac{\varpi_1 \varpi_2 (\delta)^{-\left(\frac{2}{m-1}\right)}}{\sum_{j=1}^{C} (x_i - c_j)^{-\left(\frac{2}{m-1}\right)} + (x_i - \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)} + \delta^{-\left(\frac{2}{m-1}\right)}} \quad \textbf{(3.8)}$$

$$c_j = \frac{\sum_{i=1}^{N} (\varpi_1 T_{ij})^m x_i}{\sum_{i=1}^{N} (\varpi_1 T_{ij})^m} \quad \textbf{(3.9)}$$

Parçalara ayırma (bölümleme) işlemi, amaç fonksiyonunun tekrarlamalı optimizasyonu ile gerçekleştirilmektedir. Her bir tekrarda T_{ij}, I_i, F_i üyelikleri ve c_j küme merkezleri denklem (3.6)-(3.9)'a göre güncellenmektedir. $\bar{c}_{imax}$ her bir tekrarda denklem (3.3)-(3.5) ile hesaplanmaktadır.

3.3. ÇNCO: Çekirdek Nötrozofik C-Ortalamalar Kümeleme

Doğrusal olmayan veri kümelemesinde, geleneksel kümeleme yöntemleri amaç fonksiyonu sınırlandırması nedeniyle giriş verilerini uygun kümelere ayıramaz. Dolayısıyla, bir çekirdek fonksiyonu uygulayarak giriş verilerini yüksek boyutlu bir özellik uzayına aktarmak için bir haritalama sürecine ihtiyaç vardır. Çekirdek fonksiyonu NCO algoritmasına uygulandığında, amaç fonksiyonu (3.10)'daki gibi yazılabilir:

$$\begin{aligned} J_{\text{ÇNCO}}(T, I, F, C) \\ &= \sum_{i=1}^{N} \sum_{j=1}^{C} (\varpi_1 T_{ij})^m ||\Phi(x_i) - \Phi(c_j)||^2 \\ &+ \sum_{i=1}^{N} (\varpi_2 I_i)^m ||\Phi(x_i) - \Phi(\bar{c}_{imax})||^2 + \delta^2 \sum_{i=1}^{N} (\varpi_3 F_i)^m \end{aligned} \quad \textbf{(3.10)}$$

ve

$$||\Phi(x_i) - \Phi(c_j)||^2 = K(x_i, x_i) - 2K(x_i, c_j) + K(c_j, c_j) \quad \textbf{(3.11)}$$

burada $K(x_i, x_i) = \Phi^T(x_i)\, \Phi(x_i)$ iç çarpımı göstermektedir. Eğer bir Gauss fonksiyonu çekirdek fonksiyon olarak kabul edilirse $K(x_i, x_i)$ ve $K(c_j, c_j)$ aynı olur.

Bu nedenle, yeni amaç fonksiyonu (3.12)'deki gibi olur:

$$J_{ÇNCO}(T, I, F, c)$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{C} 2\left(\varpi_1 T_{ij}\right)^m (1 - K(x_i, c_j)) + \sum_{i=1}^{N} 2(\varpi_2 I_i)^m (1 - K(x_i, \bar{c}_{imax})) + \delta^2 \sum_{i=1}^{N} (\varpi_3 F_i)^m \quad \textbf{(3.12)}$$

Bir Lagrange amaç fonksiyonu şu şekilde yapılandırılmıştır:

$$L(T, I, F, c, \lambda) = \sum_{i=1}^{N}\sum_{j=1}^{C} 2\left(\varpi_1 T_{ij}\right)^m (1 - K(x_i, c_j)) + \sum_{i=1}^{N} 2(\varpi_2 I_i)^m (1 - K(x_i, \bar{c}_{imax})) + \sum_{i=1}^{N} \delta^2 (\varpi_3 F_i)^m - \sum_{i=1}^{N} \lambda_i \left(\sum_{j=1}^{C} T_{ij} + I_i + F_i - 1\right) \quad \textbf{(3.13)}$$

Lagrange amaç fonksiyonunu minimize etmek için aşağıdaki denklemler kullanılmıştır.

$$\frac{\partial L}{\partial T_{ij}} = 2m(\varpi_1)^m \left(T_{ij}\right)^{m-1} \left(1 - K\left(x_i, c_j\right)\right) - \lambda_i \quad \textbf{(3.14)}$$

$$\frac{\partial L}{\partial I_i} = 2m(\varpi_2)^m (I_i)^{m-1} \left(1 - K(x_i, \bar{c}_{imax})\right) - \lambda_i \quad \textbf{(3.15)}$$

$$\frac{\partial L}{\partial F_i} = \delta^2 m(\varpi_3)^m (F_i)^{m-1} - \lambda_i \quad \textbf{(3.16)}$$

$$\frac{\partial L}{\partial c_j} = -2 \sum_{i=1}^{N} (\varpi_1 T_{ij})^m K'(x_i, c_j) \quad \textbf{(3.17)}$$

Norm, Öklid normu olarak belirlenmiştir. $\frac{\partial L}{\partial T_{ij}} = 0$, $\frac{\partial L}{\partial I_i} = 0$, $\frac{\partial L}{\partial F_i} = 0$ ve $\frac{\partial L}{\partial c_i} = 0$ olsun, üyelik değerleri ve küme merkezleri yeniden düzenlenirse:

$$T_{ij} = \frac{\varpi_2 \varpi_3 K(x_i, c_j)^{-\left(\frac{2}{m-1}\right)}}{\sum_{j=1}^{C} K(x_i, c_j)^{-\left(\frac{2}{m-1}\right)} + K(x_i, \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)} + \delta^{-\left(\frac{2}{m-1}\right)}} \quad \textbf{(3.18)}$$

$$I_i = \frac{\varpi_1 \varpi_3 K(x_i, \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)}}{\sum_{j=1}^{C} K(x_i, c_j)^{-\left(\frac{2}{m-1}\right)} + K(x_i, \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)} + \delta^{-\left(\frac{2}{m-1}\right)}} \quad \textbf{(3.19)}$$

$$F_i = \frac{\varpi_1 \varpi_2 (\delta)^{-\left(\frac{2}{m-1}\right)}}{\sum_{j=1}^{C} K(x_i, c_j)^{-\left(\frac{2}{m-1}\right)} + K(x_i, \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)} + \delta^{-\left(\frac{2}{m-1}\right)}} \quad \textbf{(3.20)}$$

$$c_j = \frac{\sum_{i=1}^{N} (\varpi_1 T_{ij})^m K(x_i, x_i)}{\sum_{i=1}^{N} (\varpi_1 T_{ij})^m} \quad \textbf{(3.21)}$$

Yukarıdaki denklemler, aşağıdaki adımlarda özetlenen ÇNCO algoritmasının formülasyonuna izin vermektedir.

Adım 1: $T^{(0)}$, $I^{(0)}$ ve $F^{(0)}$'ı başlat;

Adım 2: $c, m, \delta, \varepsilon, \varpi_1, \varpi_2, \varpi_3$ parametrelerini başlat;

Adım 3: Çekirdek fonksiyonunu ve parametrelerini seç;

Adım 4: Denklem (3.18)'i kullanarak k adımda $c^{(k)}$ merkez vektörlerini hesapla;

Adım 5: T_{ij}'nin en büyük ilk iki değerine ait kümelerin merkezlerini kullanarak $\bar{c}_{imax}$'ı (3.3)'teki gibi hesapla;

Adım 6: Denklem (3.15)'i kullanarak $T^{(k)}$'yı $T^{(k+1)}$'e, (3.16)'yı kullanarak $I^{(k)}$'yı $I^{(k+1)}$'e ve (3.17)'yi kullanarak $F^{(k)}$'yı $F^{(k+1)}$'e güncelle;

Adım 7: Eğer $\left|T^{(k+1)} - T^{(k)}\right| < \varepsilon$ ise dur; diğer durumlarda Adım 4'e git;

Adım 8: Her bir veriyi en büyük $TM = [T, I, F]$ değeri ile sınıflara ata:

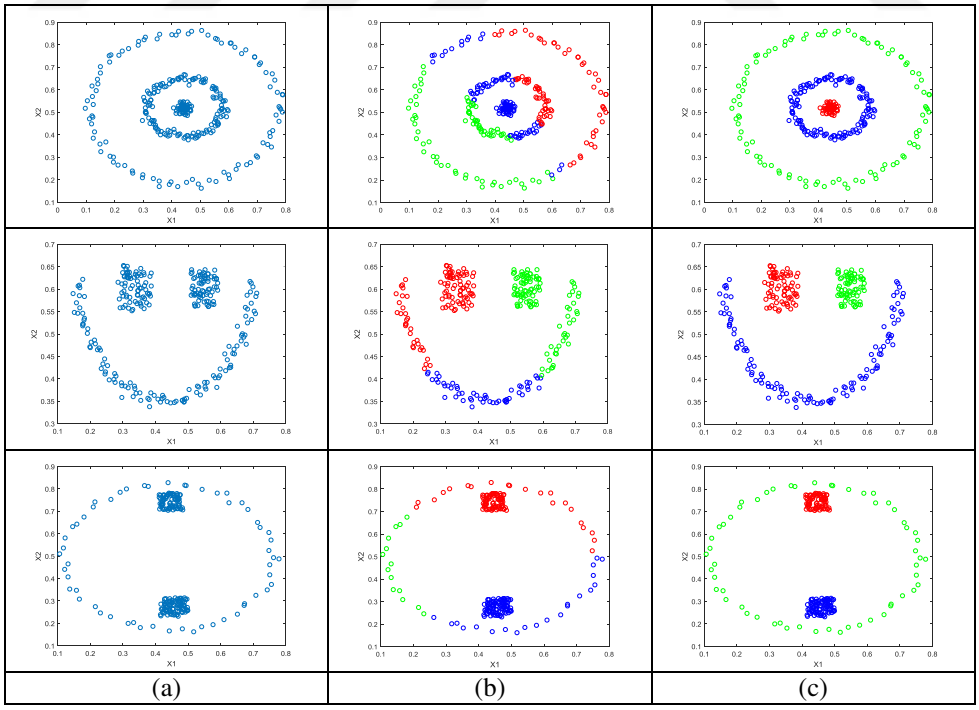
$x(i) \in k\ ncı$ sınıf eğer $k = \underset{j=1,2,\ldots,C+2}{\arg\max}\ (TM_{ij})$

3.4. Deneyler ve Sonuçlar

Bu bölümde, ÇNCO ile çekirdek BCO yöntemlerinin performanslarını karşılaştırmak amacıyla çeşitli deneyler yapılmıştır. Deneyler, yapay veri kümeleri, gerçek veri kümeleri ve görüntüler üzerinde gerçekleştirilmiştir. Hem ÇNCO hem de ÇBCO yöntemleri, $\varepsilon = 10^{-5}$ gibi aynı başlangıç parametreleriyle yürütülmüştür. Buna ek olarak, deneme yanılma yöntemi ile elde edilen ÇNCO'nın ağırlıklandırma parametreleri sırasıyla $\varpi_1 = 0{,}75$, $\varpi_2 = 0{,}125$ ve $\varpi_3 = 0{,}125$ olarak belirlenmiştir. Radyal tabanlı fonksiyon (RTF) her iki yöntem

için çekirdek fonksiyonu olarak kabul edilmiştir. ÇBCO için RTF çekirdeğinin parametresi aralık arama (interval search) yöntemi ile elde edilmiştir. Belirli bir aralık için, ÇBCO kümeleme algoritması uygun bir artırım değeriyle çalıştırılmış ve kümeleme hatasının en aza indirildiği yerde optimum RTF parametresi seçilmiştir. Aynı işlem ÇNCO için de kabul edilmiştir. ÇNCO yönteminde, RTF çekirdek parametresi ve delta gibi iki ayarlanabilir parametreye sahip olduğundan uygun artırım değeri ile 2 boyutlu (2B) bir aralık arama düşünülmüştür. Deneyler, 32 GB bellekli ve Intel Core i7-4810 işlemcili bir bilgisayarda MATLAB 2014b yazılımı kullanılarak gerçekleştirilmiştir.

Deneysel çalışmalar, NCO kümeleme üzerinde çekirdek fikrinin etkisini göstermek amacıyla NCO ve ÇNCO arasında bir karşılaştırma ile başlamıştır. Açık bir şekilde belirtmek gerekirse KO, BCO ve NCO tipi kümeleme algoritmaları, doğrusal olmayan veri kümelerini uygun bir şekilde kümeleyemez. Bu etkiyi göstermek için Şekil 3.1'de gösterildiği gibi üç farklı şekillendirilmiş veri kümesi içeren doğrusal olmayan çeşitli yapay veri kümeleri kullanılmıştır. Şekil 3.1'de (a), (b) ve (c) sütunları sırasıyla ham veriyi, NCO ve ÇNCO sonuçlarını göstermektedir. Sonuçlara göre, ÇNCO kesin-referans (ground-truth) kümeleme sonuçlarını elde ederken, NCO kesin kümeleri elde edememiştir.



Şekil 3.1. NCO ve ÇNCO'nun çeşitli yapay veri kümeleri üzerinde karşılaştırılması (a) Ham veri, (b) NCO sonuçları, (c) ÇNCO sonuçları

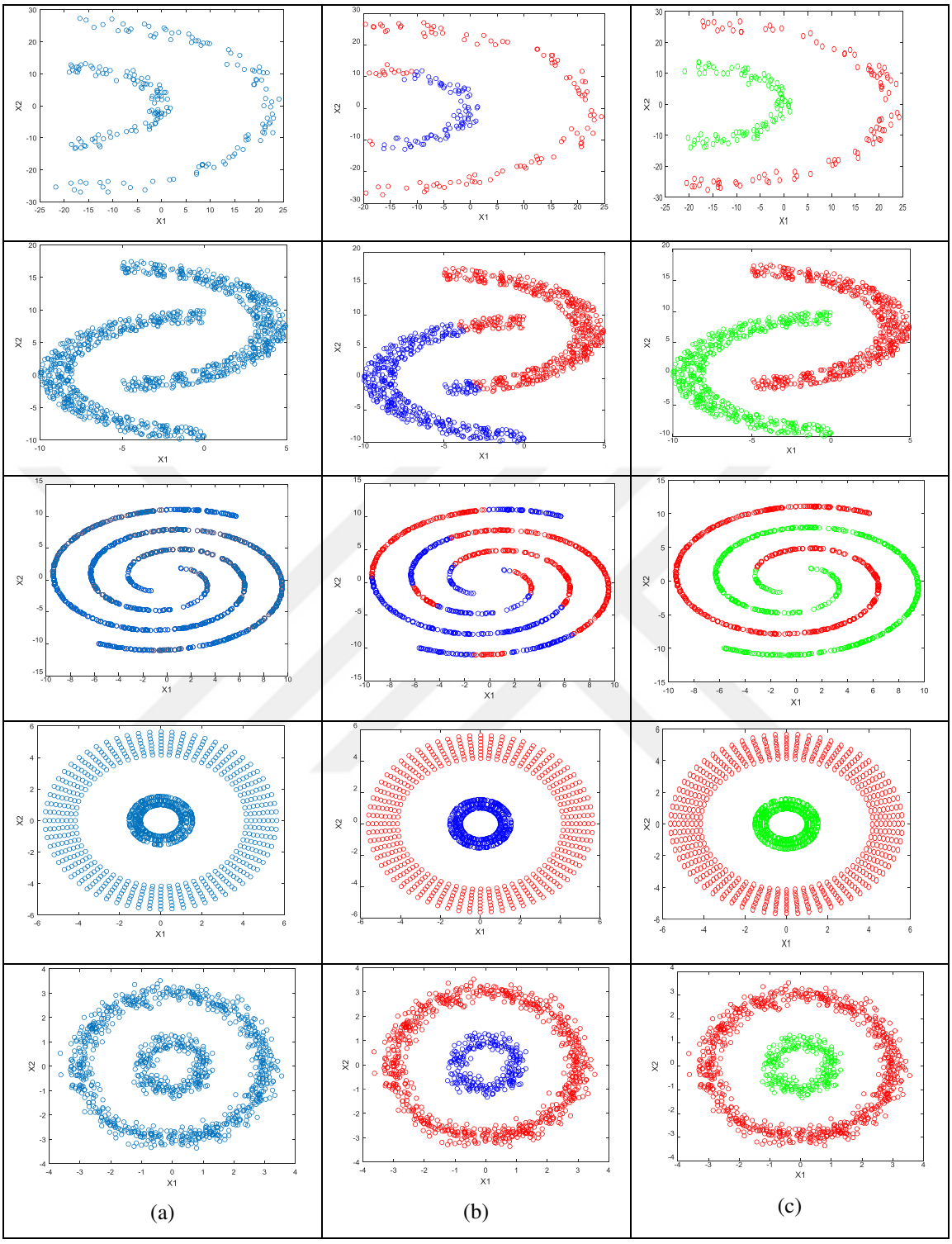
3.4.1. Yapay Veri Örneği 1

İlk deney türünde, Şekil 3.2'de gösterildiği gibi iki-sınıflı (kümeli) veri kümesi ele alınmıştır. Şekil 3.2'nin ilk sütununda veri kümelerinin ham hali bulunmaktadır. İkinci sütunda çekirdek BCO'nun kümeleme sonuçları verilmiştir ve üçüncü sütunda da ÇNCO sonuçları gösterilmiştir. Şekil 3.2'nin ilk satırında, 200 örnekten oluşan "*Two-kernel*" veri kümesi gösterilmiştir. Her bir sınıf 100 örneğe sahiptir. Arama işlemi ÇBCO'nın RTF parametresini otomatik olarak 223 seçmiş ve benzer şekilde RTF parametresine 280, delta parametresine 100 değeri atanmıştır. Elde edilen kümeler farklı renklerde gösterilmiştir. Şekil 3.2 ve Şekil 3.3'te görülebileceği gibi, ÇBCO yöntemi geçerli kümeleri üretememiştir. Bazı örnekler hatalı bir şekilde kümelenmiştir, özellikle bazı içteki örnekler hatalı kümelenmiştir. Diğer taraftan, ÇNCO her iki kümeyi hatasız olarak kümelemiştir. ÇNCO birkaç kez çalıştırıldığında her zaman doğru kümeler ürettiğini de belirtmek gerekmektedir.

Bir başka örnek Şekil 3.2'nin ikinci satırında gösterilmiştir. Bu veri kümesine "*Half-Moons*" denmektedir. Doğrusal olmayan, ayrılabilir iki yarım ay şeklindeki veri kümesi 1000 örnek içermektedir. ÇBCO için RTF parametresi 9,3 olarak elde edilmiş ve ÇNCO için RTF çekirdek parametresi 2,5 ve delta değeri 10 olarak atanmıştır. ÇBCO bazı veri örneklerini doğru bir şekilde sınıflandıramamıştır. Özellikle yarım ayların iç kuyruk bölümündeki veri noktaları yanlış sınıflandırılmıştır. Önceki veri kümesinde olduğu gibi, ÇNCO kümelemeyi hatasız olarak gerçekleştirmiş ve tüm veri noktaları doğru kümelere atanmıştır.

Deneylerde diğer bir popüler doğrusal olmayan veri kümesi "*Two-Spirals*" düşünülmüştür. Ham veri kümesi Şekil 3.2'nin üçüncü satırında görülmektedir. RTF çekirdek parametresi hem ÇBCO hem de ÇNCO için arama algoritması tarafından 1,7 olarak üretilmiştir ve delta değeri 7,0'dir. Kümeleme sonuçları Şekil 3.2'nin üçüncü satırında gösterildiği üzere her iki yöntem de doğru kümeleri üretememiştir. Özellikle, ÇBCO yöntemi anlamsız kümeler üretmiş ve her iki kümede de yanlış sınıflandırma yapmıştır. Öte yandan, ÇNCO daha mantıklı sonuçlar üretmiş ve kümelerden biri doğru olarak sınıflandırılmıştır (yeşil renkli halkalar). Ayrıca diğer kümenin (kırmızı renkli halkalar) yalnızca bir kısmı yanlış sınıflandırılmıştır.

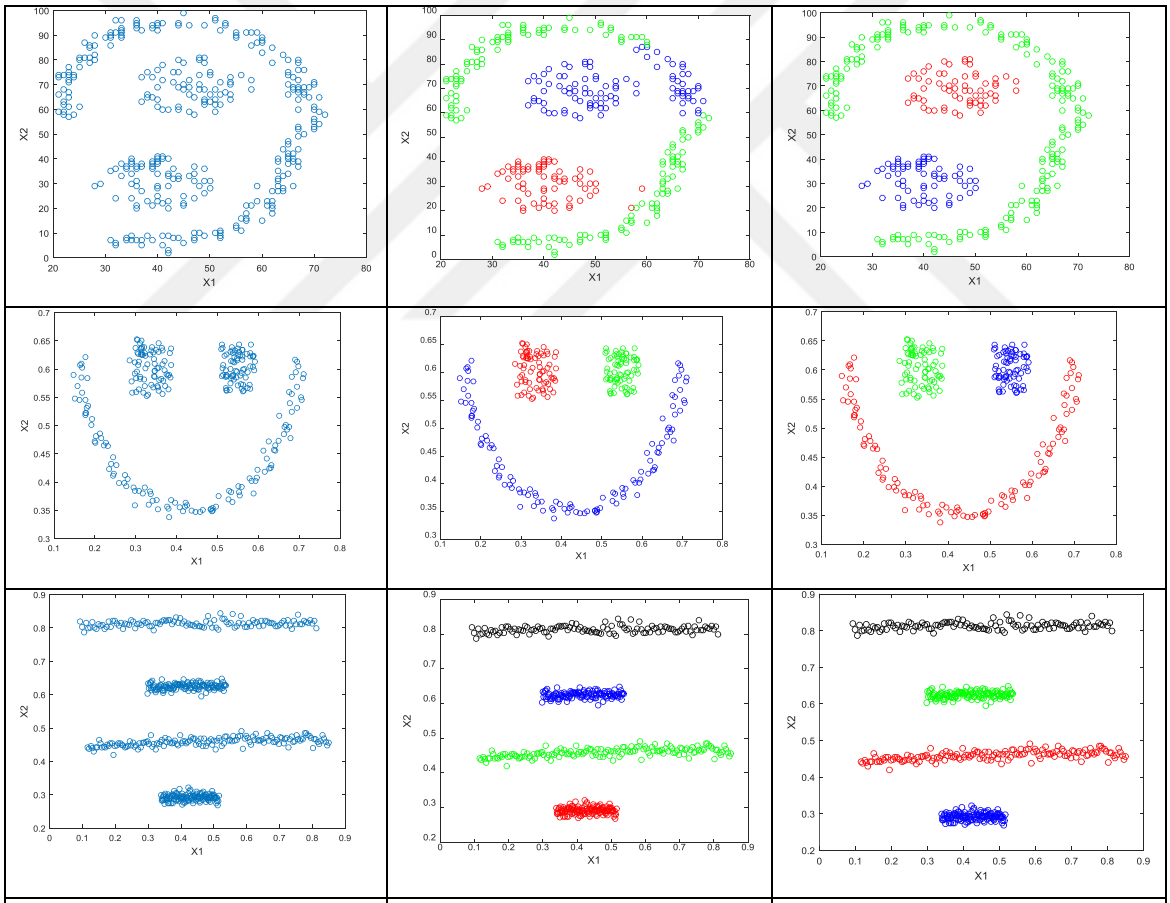
Şekil 3.2'nin son iki satırında benzer veri kümeleri ile deneyler gerçekleştirilmiştir. Her iki durumda da hem ÇBCO'nun hem de ÇNCO'nun net olarak ürettiği dairesel kümeler vardır. Her iki yöntemin ulaştığı kümeleme doğruluğu %100'dür.



Şekil 3.2. İki-sınıflı yapay veri kümelerinde ÇBCO ile ÇNCO'nun karşılaştırılması (a) Ham veri, (b) ÇBCO sonuçları, (c) ÇNCO sonuçları

3.4.2. Yapay Veri Örneği 2

Küme sayısının ikiden fazla olduğu yapay veri kümeleri üzerinde daha ileri deneyler yapılmıştır. İlgili sonuçlar Şekil 3.3'te gösterilmiştir. Şekil 3.2'ye benzer şekilde, Şekil 3.3'teki ilk sütun ham yapay veri kümelerini göstermektedir. Şekil 3.3'ün ikinci sütununda ÇBCO kümeleme sonuçları ve üçüncü sütununda ise önerilen ÇNCO kümeleme sonuçları bulunmaktadır. Üç veya dört küme içeren yapay veri kümeleri deneylerde ele alınmıştır. Elde edilen sonuçlar değerlendirildiğinde, "*Ear*" veri kümesi dışında her iki yöntemde tüm yapay veri kümeleri için %100 doğru kümeleme sonucu ürettiği görülmüştür. "*Ear*" veri kümesi için ÇNCO yöntemi, ÇBCO yönteminden daha doğru sonuçlar üretmiştir. Çekirdek parametresi ve delta değeri ilk deneydekine benzer şekilde ayarlanmıştır.



Şekil 3.3. ÇBCO ile ÇNCO'nun diğer yapay veri kümelerinde karşılaştırılması (a) Ham veri, (b) ÇBCO sonuçları, (c) ÇNCO sonuçları

3.4.3. Spektral Kümeleme Yöntemleriyle Karşılaştırma

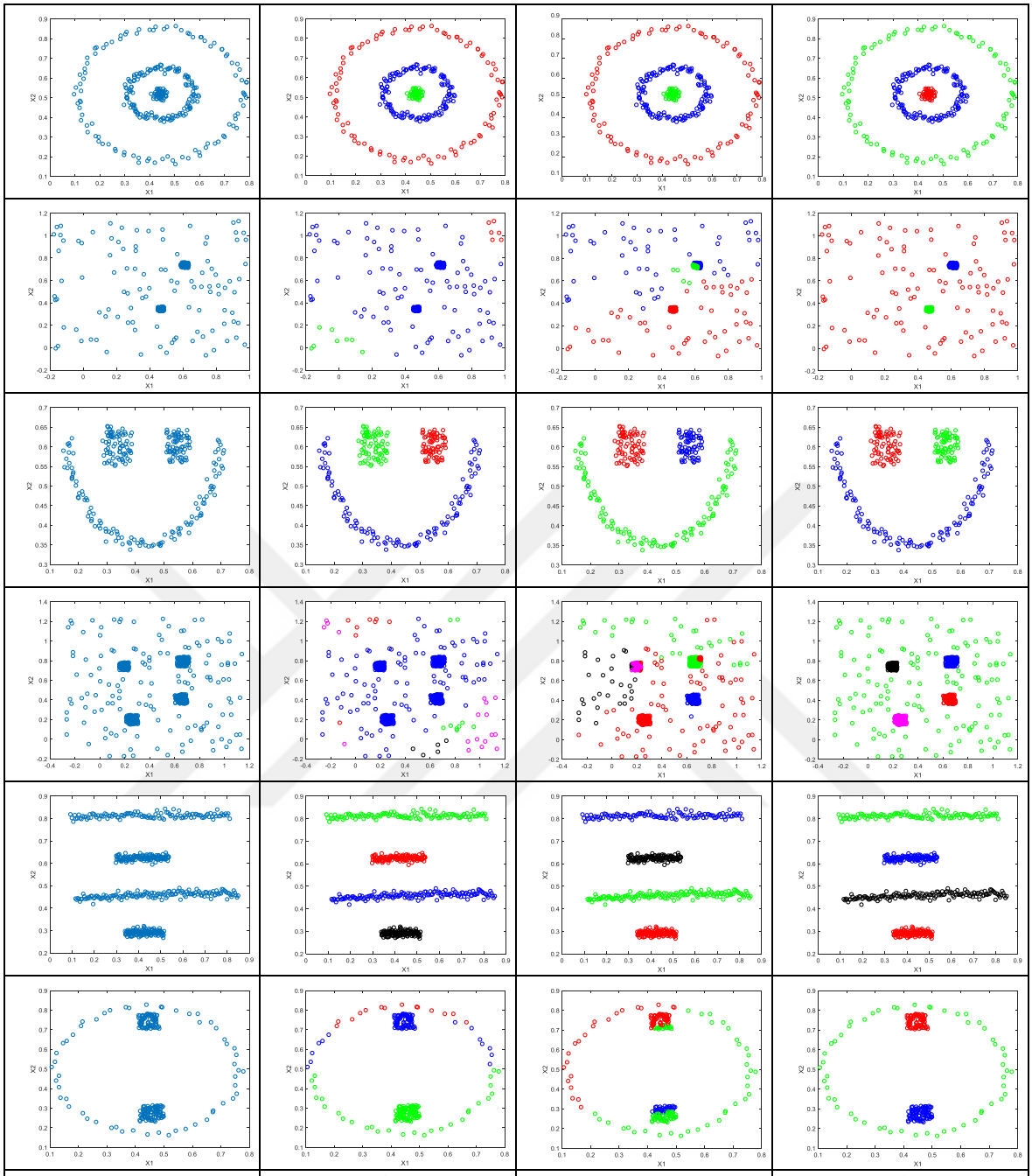
ÇNCO yöntemi ayrıca, Spektral Kümeleme (SK) [104] ve Spektral Çoklu Manifold (katman) Kümeleme (SÇMK) [105] adı verilen özdeğer tabanlı iki kümeleme yöntemiyle karşılaştırılmıştır.

SK, giriş verisinin benzerlik matrisinin özdeğerlerini kullanan popüler kümeleme algoritmasıdır. Özdeğerleri kullanmasının temel amacı, kümeleme öncesi boyutsal küçülme gerçekleştirmektir. Son olarak, küçültülmüş veri kümesini kümelemek için KO algoritmasını kullanmaktadır.

SÇMK algoritması aynı zamanda, çoklu düzgün düşük boyutlu manifoldları (katman) SK algoritmasına bütünleştirerek onun başarımını iyileştiren bir SK tabanlı kümeleme yöntemidir. Daha sonra, SÇMK uygun bir yakınlık matrisi oluşturmak için örneklenen verilerin yerel geometrik bilgilerini kullanmaktadır. Son olarak, SK bu yakınma matrisi ile verileri gruplamak için kullanılmaktadır.

Karşılaştırmalar, Şekil 3.4'te gösterilen 6 yapay veri kümesi üzerinde yapılmıştır. İlk sütunda ham yapay veri kümeleri gösterilirken, ikinci, üçüncü ve dördüncü sütunlarda sırasıyla SK, SÇMK ve ÇNCO gösterilmektedir. Şekil 3.4'teki birinci, üçüncü ve beşinci satırlardan görülebileceği üzere, tüm yöntemler yapay veri kümeleri için kesin-referans (ground-truth) üretmiştir. Öte yandan, sadece ÇNCO yöntemi, Şekil 3.4'ün ikinci, dördüncü ve altıncı satırlarında görüldüğü gibi yapay veri kümeleri için kesin-referans (ground-truth) doğru kümeleme sonuçları vermiştir. Bu sonuçlar hem SK hem de SÇMK yöntemlerinin gürültü kümesi içeren veri kümelerini kümeleyemediklerini göstermektedir. Başka bir deyişle, ÇNCO doğrusal olmayan veri kümeleri üzerinde benzer bir başarım yakalamaktadır. Buna ek olarak, ÇNCO gürültü kümesi içeren verilerde daha iyi başarım sergilemektedir.

Şekil 3.4. SK, SÇMK ve ÇNCO yöntemlerinin 6 yapay veri kümesinde karşılaştırılması (a) Ham veri, (b) SK sonuçları, (c) SÇMK sonuçları, (d) ÇNCO sonuçları

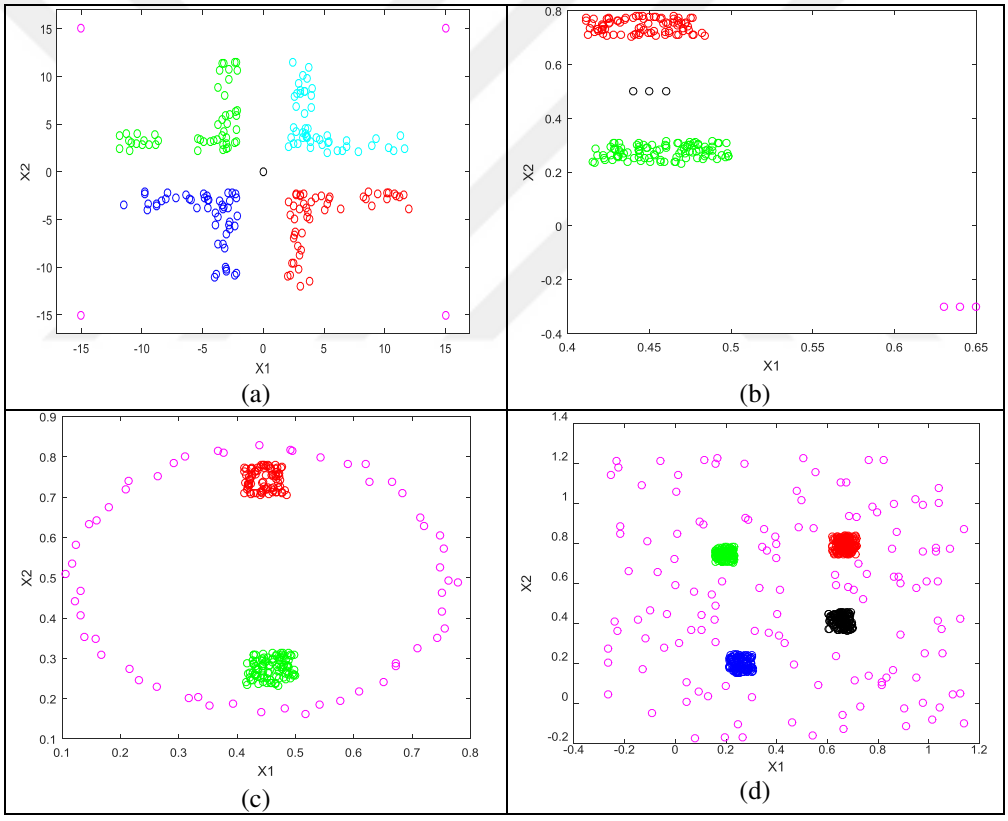


(a) (b) (c) (d)

3.4.4. Gürültülü ve Aykırı Yapay Veri Kümeleri

Daha önce belirtildiği gibi NCO algoritması, gürültülü ve aykırı veri noktalarının kümelenmesinde daha iyi başarım göstermiştir. ÇNCO algoritmasının gürültülü ve aykırı veri noktaları içeren veri kümeleri ile iyi çalıştığını göstermek için çeşitli yapay veri kümeleri üzerinde birçok deney yapılmıştır. Elde edilen sonuçlar Şekil 3.5'te gösterilmiştir. Şekil 3.5 (a)'da doğrusal olarak ayrılabilen dört küme içeren "*Corner*" veri kümesi bulunmaktadır. Dört kümenin ortasındaki bir veri noktası yapay olarak konumlandırılmıştır.

Ayrıca Şekil 3.5 (a)'da gösterildiği üzere dört veri noktası daha yapay olarak yerleştirilmiştir. ÇNCO, "*Corner*" veri kümesinde sadece dört kümenin hepsini doğru olarak kümelemekle kalmamış, aynı zamanda gürültülü ve aykırı veri noktalarını tespit etmiştir. Siyah ve macenta renkleri sırasıyla gürültüyü ve aykırı veri noktalarını temsil etmektedir. Benzer bir senaryo, Şekil 3.5 (b)'de gösterildiği gibi iki kümeli bir durum için oluşturulmuş ve bu senaryo için benzer başarılı kümeleme sonuçları da elde edilmiştir. Şekil 3.5 (c)'de doğrusal olarak ayrılabilir iki kümeyi çevreleyen dairesel aykırı veri noktaları ele alınmıştır. Bunlara ek olarak, Şekil 3.5 (d)'de çok fazla aykırı değere sahip başka bir yapay veri kümesi gösterilmiştir. Şekil 3.5 (c) ve (d)'deki bu iki veri kümesi için, önerilen ÇNCO algoritması makul kümeler bulmuştur.



Şekil 3.5. Gürültülü ve aykırı veri noktalarında ÇNCO kümeleme sonuçları

3.4.5. Farklı Çekirdeklerin Çeşitli Yapay Veri Kümelerinde Kullanılması

RTF çekirdek tabanlı ÇNCO daha iyi sonuçlar vermesine rağmen, bu başlık altında, ÇNCO'nun çeşitli yapay veri kümeleri için çeşitli çekirdeklerle daha iyi sonuçlar ürettiği gösterilmiştir. Bu amaçla sırasıyla "*çoklu*" (poly), "*dalga*" (wave) ve "*doğrusal*" (linear) çekirdekler kullanılmıştır. Deneylerde kullanılan çekirdek fonksiyonlarının (RTF, çoklu, dalga, doğrusal) tanımı;

$$K_{RTF}(x, x_i) = e^{(-\frac{\|x-x_i\|_2^2}{\sigma^2})} \tag{3.22}$$

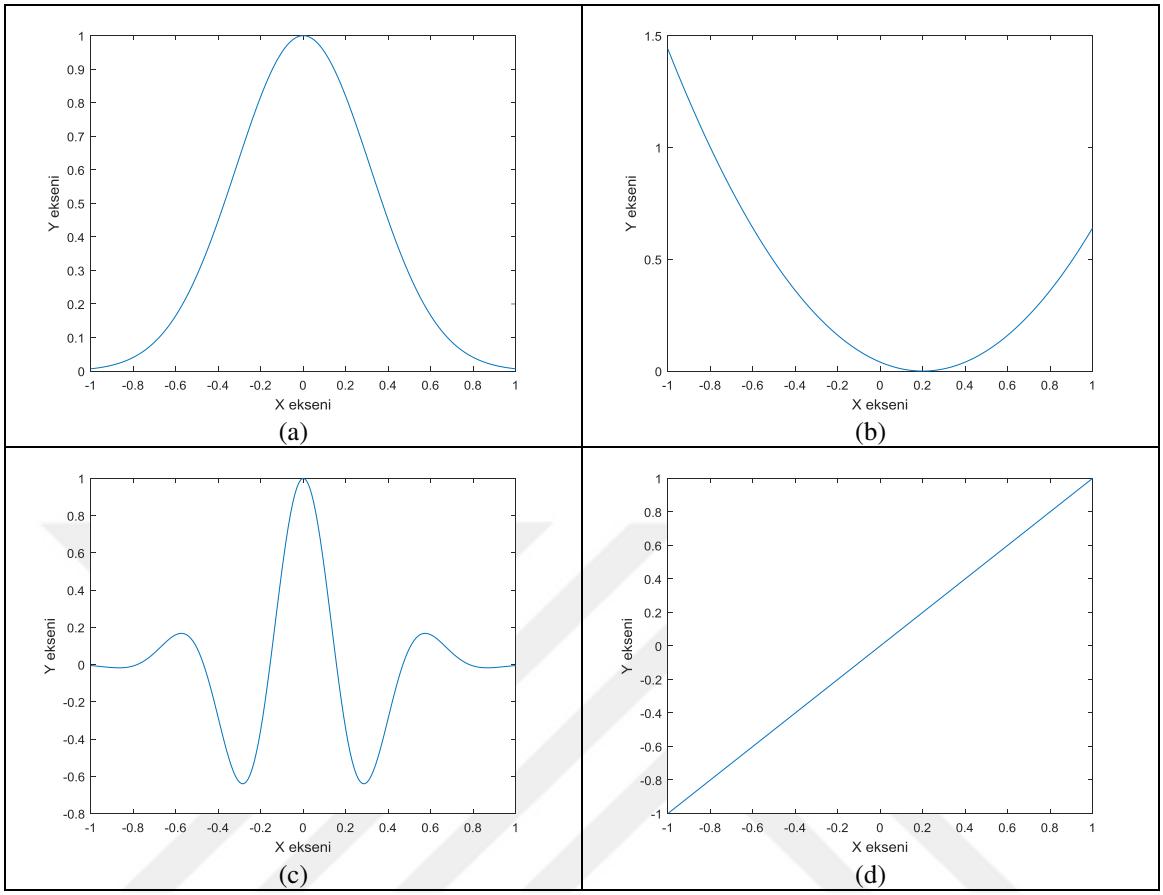
$$K_{coklu}(x, x_i) = (x_i^T x + \tau)^d \tag{3.23}$$

$$K_{dalga}(x, x_i) = \cos(\frac{\alpha(x - x_i)}{\beta}) e^{(-\frac{\|x-x_i\|_2^2}{\sigma^2})} \tag{3.24}$$

$$K_{dogrusal}(x, x_i) = x_i^T x \tag{3.25}$$

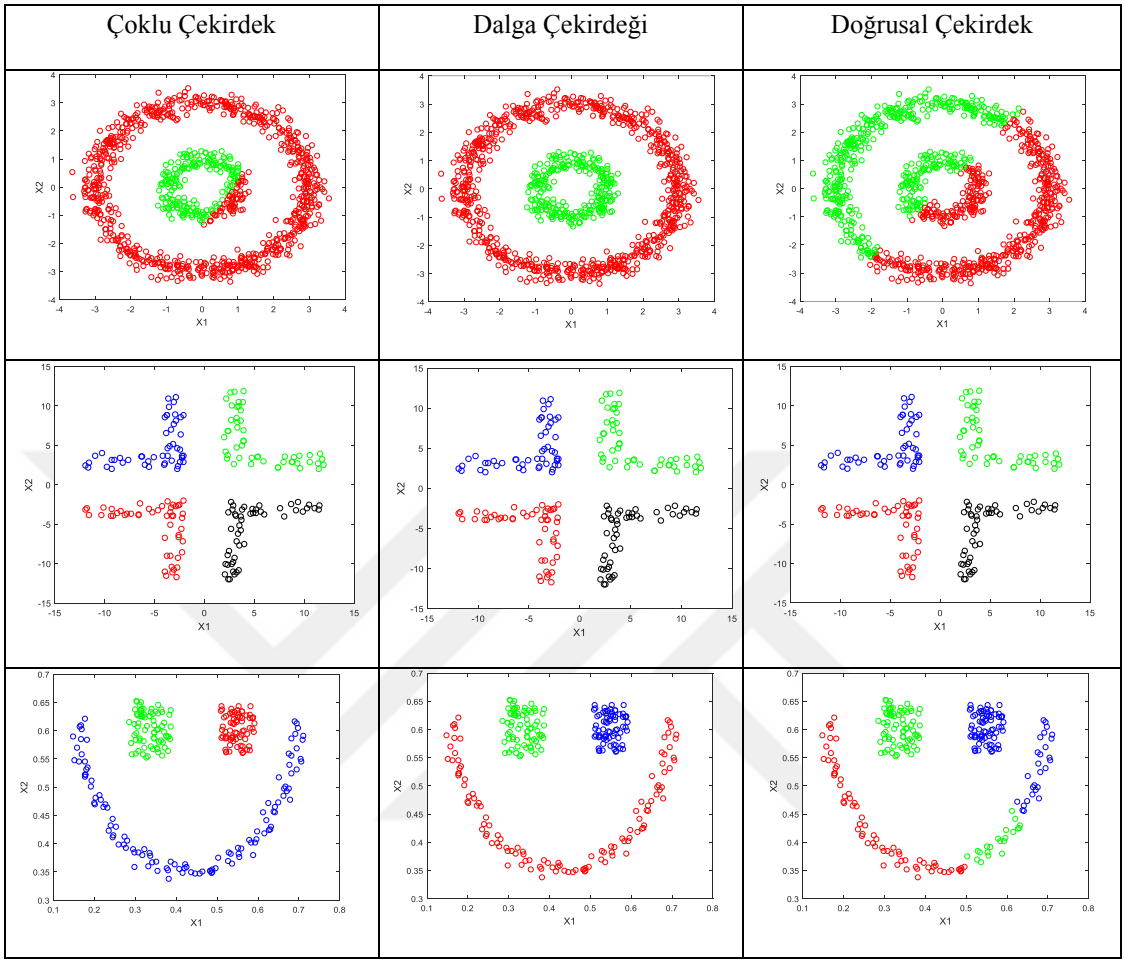
Denklem (3.22)'den görülebileceği üzere, RTF çekirdeği yalnızca bir ayarlanabilir parametreye sahiptir. Denklem (3.25) "*doğrusal*" çekirdeğin herhangi bir ayarlanabilir parametreye sahip olmadığı gösterirken, "*çoklu*" ve "*dalga*" çekirdekler sırasıyla iki ve üç ayarlanabilir parametreye sahiptirler.

Denklem (3.22)-(3.25)'te tanımları verilen çekirdek fonksiyonları x ekseninde -1 ile 1 arasında 0,01 adım aralığı kullanılarak Şekil 3.6'da çizdirilmiştir. Şekil 3.6 (a)-(d), sırasıyla RTF çekirdeğini, "*çoklu*" çekirdeği, "*dalga*" çekirdeği ve "*doğrusal*" çekirdeği göstermektedir.



Şekil 3.6. Deneylerde kullanılan **ç**ekirdek fonksiyonları (a) RTF, (b) Çoklu, (c) Dalga, (d) Doğrusal

Bu çekirdeklerin parametreleri de ayarlanmıştır. Birkaç kümeleme sonucu Şekil 3.7'de gösterilmiştir. "*dalga*" çekirdeği ile elde edilen sonuçlar oldukça başarılı sayılmıştır. "*çoklu*" çekirdekler de sadece birkaç veri noktasını yanlış kümeleyerek daha iyi sonuçlar vermiştir. "*doğrusal*" çekirdek, sadece doğrusal olarak ayrılabilen "*Corner*" veri kümesini doğru kümeleyerek en kötü sonucu elde etmiştir. "*doğrusal*" çekirdek, dairesel bir veri yapısına sahip veri kümelerini doğru şekilde ayıramamıştır.



Şekil 3.7. Farklı çekirdek fonksiyonları ile ÇNCO kümeleme sonuçları

3.4.6. Gerçek Veri Kümesi Örneği 1

Elde edilen kümeleme sonuçlarını her iki yöntemle de karşılaştırmak ve değerlendirmek için çeşitli gerçek veri kümeleri kullanılmıştır. Bu amaçla her şeyden önce araştırmacılar tarafından yaygın olarak kullanılan ve popüler olan "*Iris*" veri kümesi ele alınmıştır. "*Iris*" veri kümesinde, isimleri Iris Setosa, Iris Versicolor, Iris Virginica olan ve her biri 50 örnekten oluşan üç çeşit Iris çiçeği vardır. Her örnekte çanak yaprak uzunluğu, çanak yaprak genişliği, taç yaprak uzunluğu ve taç yaprak genişliği olmak üzere dört öznitelik bulunmaktadır. Üç kümeden biri diğer ikisinden açıkça ayrılırken, bu iki sınıf bazı çakışmaları kabul etmektedir. RTF çekirdek parametresi 0,001 ve delta 3'tür. Tablo 3.1'den görülebileceği gibi, Setosa %100 doğru kümeleme oranı ile kümelenmiş ancak Versicolor

ve Virginica gibi diğer kümeler tam olarak kümelenmemiştir. Dört Virginica örneği yanlışlıkla Versicolor olarak ve bir Versicolor örneği yanlışlıkla Virginica olarak kümelenmiştir. Toplam doğruluk %96,67'dir.

Tablo 3.1. "*Iris*" veri kümesinin ÇNCO kümeleme sonuçları

		Gerçek		
		Setosa	**Versicolor**	**Virginica**
Kümeleme	**Setosa**	50	0	0
	Versicolor	0	49	4
	Virginica	0	1	46

3.4.7. Gerçek Veri Kümesi Örneği 2

İkinci gerçek veri kümesindeki deneyler "*Wine*" veri kümesi üzerinde yürütülmüştür. "*Wine*" veri kümesi İtalya'nın aynı bölgesinde üretilen şarapların kimyasal analiz sonuçlarına dayanılarak oluşturulmuştur. Veri kümesi 13 öznitelik ve 3 küme içermektedir. Toplam örnek sayısı 178'dir, her bir küme sırasıyla 59, 71 ve 48 örneğe sahiptir. Yapılan deneysel çalışmalarda, RTF çekirdek parametresi 0,0095, delta 100 olarak seçilmiş ve elde edilen sonuçlar Tablo 3.2'de verilmiştir; burada kümelerden hiçbiri %100 doğrulukta sınıflandırılmamıştır ve yanlış sınıflandırılmış örneklerin toplam sayısı 16'dır. Böylece toplam doğruluk %91,01 olmuştur.

Tablo 3.2. "*Wine*" veri kümesinin ÇNCO kümeleme sonuçları

		Gerçek		
		Sınıf 1	**Sınıf 2**	**Sınıf 3**
Kümeleme	**Sınıf 1**	57	7	0
	Sınıf 2	2	58	1
	Sınıf 3	0	6	47

3.4.8. Gerçek Veri Kümesi Örneği 3

Gerçek veri kümesiyle ilgili üçüncü deney, "*Parkinson*" veri kümesi üzerinde gerçekleştirilmiştir. Veri kümesi 22 öznitelik içermektedir ve toplam örnek sayısı 195'tir. Bunların 147'si Parkinson ve geri kalan 48'i sağlıklıdır. Veriler, sağlıklı insanları Parkinson

hastalığı (PH) olanlardan ayırmada kullanılmaktadır. Başka bir deyişle, iki küme vardır. Yapılan çalışmalarda, RTF çekirdek parametresi 0,005, delta 10 olarak seçilmiştir. Tablo 3.3'te görüldüğü üzere 29 PH sağlıklı olarak kümelenmiş ve 4 sağlıklı örnek PH olarak kümelenmiştir. Dolayısıyla toplam doğruluk %83,08'dir.

Tablo 3.3. "*Parkinson*" veri kümesinin ÇNCO kümeleme sonuçları

		Gerçek	
		Hasta	**Sağlıklı**
Kümeleme	**Hasta**	118	4
	Sağlıklı	29	44

Aynı deneyler ÇBCO ile de yapılmış ve her iki yöntemin doğruluk karşılaştırmaları Tablo 3.4'te verilmiştir. Önerilen ÇNCO'nin tüm gerçek veri kümeleri için ÇBCO'dan daha iyi sonuç verdiği apaçık ortadadır.

Tablo 3.4. Her iki yöntemin gerçek veri kümeleri üzerindeki başarım karşılaştırmaları

	ÇBCO	**ÇNCO (önerilen yöntem)**
Iris	%95,33	%96,67
Wine	%89,89	%91,01
Parkinson	%80,00	%83,08

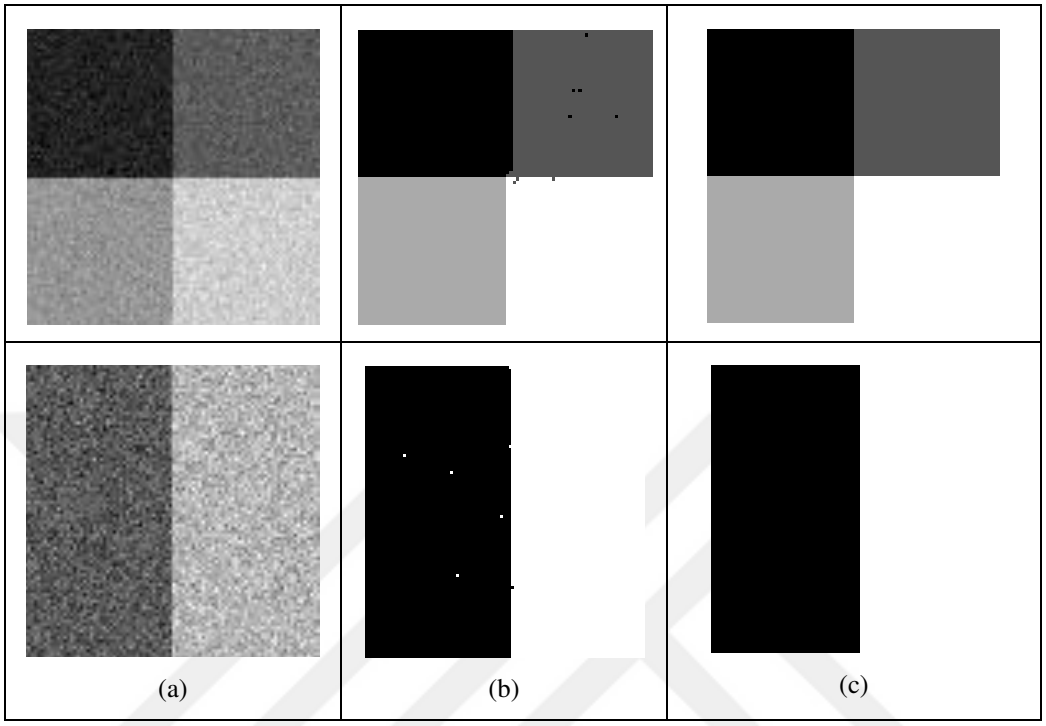
3.4.9. Görüntü Bölütleme Örneği 1

Görüntü bölütleme deneylerini tanımlamadan önce her iki çekirdek yöntemi için uzamsal bilginin kullanıldığını belirtmek gerekir. Uzamsal bilgi [106] hem ÇNCO hem de ÇBCO için düşünülmüştür.

İki farklı sentezlenmiş görüntü kullanılmıştır. İlkinde 100×100 boyutlu dört sınıf bulunmaktadır. İlgili gri seviye değerleri sırasıyla 50 (sol üst), 100 (sağ üst), 150 (sol alt) ve 200'dür (sağ alt). Görüntüye, $\mu = 0$ ve $\sigma = 25$ Gauss gürültüsü eklenmiştir. İkinci görüntünün iki kümesi vardır. İlgili gri seviye değerleri sol sütun için 30 ve sağ sütun için 80'dir. Görüntü, $\mu = 0$ ve $\sigma = 15$ Gauss gürültüsüyle aşındırılmıştır.

Elde edilen sonuçlar Şekil 3.8'de verilmiştir. İlk sütun (a)'da gürültülü görüntü, ikinci sütun (b)'de ÇBCO sonuçları ve üçüncü sütun (c)'de ÇNCO sonuçları gösterilmiştir. Görsel

inceleme yapıldığında önerilen yöntemin kesin bölümler ürettiği ve ÇBCO yönteminin her iki görüntü için birkaç yanlış sınıflandırılmış piksel ürettiği görülmüştür.

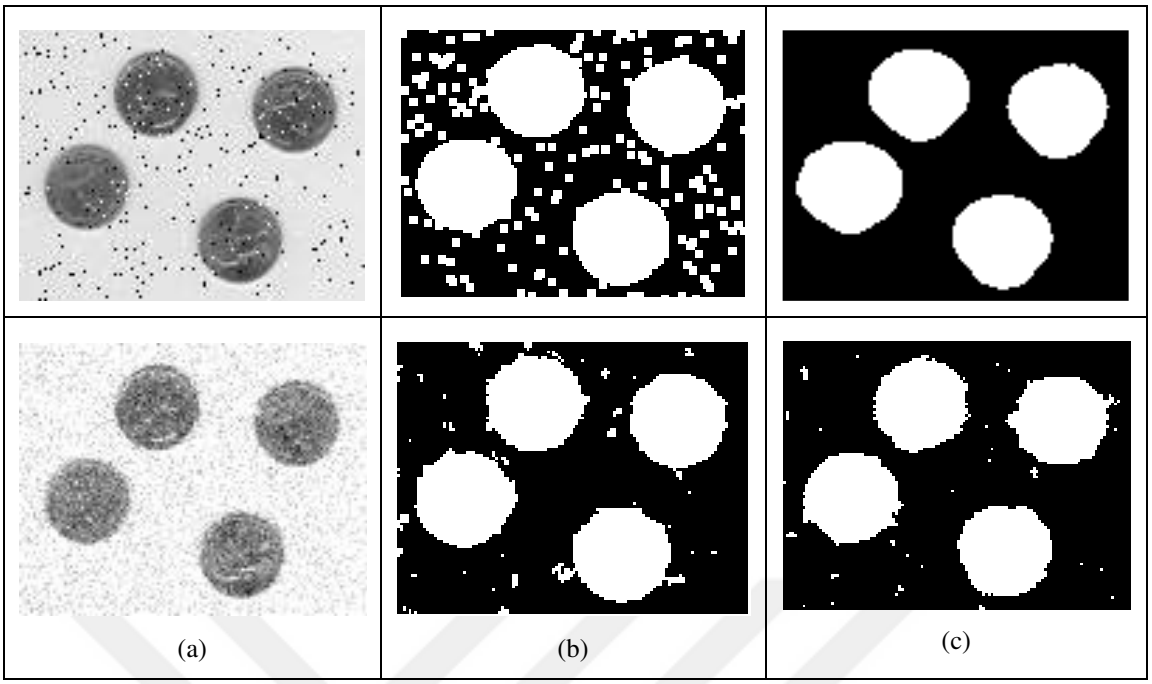


Şekil 3.8. ÇBCO ve ÇNCO'nin görüntü bölütleme uygulamasında karşılaştırılması (a) Gürültülü görüntü, (b) ÇBCO kümeleme sonuçları, (c) ÇNCO kümeleme sonuçları

3.4.10. Görüntü Bölütleme Örneği 2

İkinci görüntü bölütleme deneyinde, "*Eight*" görüntüsü farklı gürültü türleri ve seviyeleri ile kullanılmıştır. Gürültü türleri "*Tuz ve Biber*" (Salt and Pepper) ve "*Gauss*" (Gaussian)'dır. "*Tuz ve Biber*" için gürültü yoğunluğu 0,04 ve "*Gauss*" gürültü parametrelerinin ortalaması 0, varyansı 0,1'dir.

Gerçek bir görüntü için elde edilen sonuçlar Şekil 3.9'da verilmiştir. Görsel inceleme sonrasında, önerilen ÇNCO'nun hem gürültü türlerinde hem de seviyelerinde ÇBCO'dan daha iyi kümeleme sonuçları üretmiştir.



Şekil 3.9. ÇBCO ve ÇNCO'nin görüntü bölütleme uygulamasında karşılaştırılması (a) Gürültülü görüntü, (b) ÇBCO kümeleme sonuçları, (c) ÇNCO kümeleme sonuçları

4. NÖTROZOFİK AĞIRLIKLANDIRILMIŞ AŞIRI ÖĞRENME MAKİNESİ

4.1. Giriş

Önceki bölümde, veri kümelenmesi için bir kümeleme algoritması olan NCO'dan bahsedilmişti [13, 14]. Bu bölümde, son derece dengesiz veri kümelerindeki AÖM'nin dezavantajlarının üstesinden gelmek için NCO kümeleme algoritmasını kullanan yeni bir Nötrozofik Ağırlıklandırılmış Aşırı Öğrenme Makinesi (NAAÖM) yöntemi önerilmiştir. Bir veri kümesindeki gürültüler ve aykırı veri noktaları bir tür belirsizlik olarak değerlendirilebilir. Nötrozofik küme (NK), belirsiz bilgi işleme için başarılı bir şekilde uygulanmaktadır. NK, verilerin belirsizlik bilgileriyle başa çıkması avantajını göstermektedir ve halen veri analizi ve sınıflandırma uygulaması için desteklenen bir tekniktir. NAAÖM yönteminde, NCO bir örneğin aitlik, gürültü ve belirsizlik üyeliklerini belirlemek için görevlendirilmiş ve daha sonra bu veri örneğinin ağırlık değerinin hesaplanmasında kullanılmıştır [13, 14, 107].

Önerilen NAAÖM yöntemi, NCO'daki ağırlıklar kullanılarak ve AÖM temel alınarak tasarlanmıştır ve dengesiz veri kümesi sınıflandırması için uygun yapıdadır. Yöntemin değerlendirilmesi amacıyla yapay veri kümeleri ve gerçek veri kümeleri üzerinde sınıflandırma uygulamaları yapılmıştır. Elde edilen sonuçlar, AÖM, AAÖM ve DVM yöntemleriyle karşılaştırılmış ve önerilen NAAÖM yönteminin dengesiz veri kümelerini sınıflandırma uygulamalarında bu yöntemlerden daha etkili olduğu görülmüştür.

4.2. Önerilen Yöntem

4.2.1. Aşırı Öğrenme Makinesi

Eğim tabanlı bir öğrenme yöntemi olarak bilinen geriye yayılım, yavaş yakınsama hızından kötü etkilenmektedir. Bununla birlikte, yerel minimuma takılması eğim tabanlı öğrenme algoritmasının diğer bir dezavantajı olarak gösterilebilir. AÖM, eğim tabanlı öğrenme yöntemlerinin eksikliklerinin üstesinden gelen alternatif bir yöntem olarak Huang ve arkadaşları tarafından önerilmiştir [16]. AÖM, giriş ağırlıklarının ve gizli eşik değerlerinin rastgele seçildiği bir tek-gizli katmanlı ileri beslemeli (TGKİB) ağ olarak tasarlanmıştır. Bu ağırlıkların, eğitim süreci boyunca ayarlanması gerekmemektedir. Çıkış

ağırlıkları, gizli katman çıkış matrisinin Moore-Penrose genelleştirilmiş tersiyle analitik olarak belirlenmektedir.

Matematiksel olarak, L gizli düğümleri ve $g(.)$ aktivasyon fonksiyonu olmak üzere AÖM'nin çıkışı şu şekilde yazılabilir;

$$o_i = \sum_{j=1}^{L} \beta_j g(a_j, b_j, x_i), \quad i = 1,2, \dots, N \tag{4.1}$$

burada o_i , i 'inci çıkış düğümü, $\beta_j = [\beta_{j1}, \beta_{j2}, \dots, \beta_{jN}]^T$ çıkış ağırlık vektörü, $a_j = [a_{j1}, a_{j2}, \dots, a_{jN}]^T$ ağırlık vektörü, b_j, j'ninci gizli düğümün eşik değeri, x_i, i'inci giriş verisi ve N örnek sayısıdır. AÖM bu N tane örneği sıfır hata ile öğrenirse, (4.1) aşağıdaki şekilde güncellenebilir;

$$t_i = \sum_{j=1}^{L} \beta_j g(a_j, b_j, x_i), \quad i = 1,2, \dots, N \tag{4.2}$$

burada t_i gerçek çıkış vektörünü göstermektedir. Denklem (4.2), (4.3)'te gösterildiği gibi kısa ve öz bir şekilde yazılabilir;

$$H\beta = T \tag{4.3}$$

burada $H = \{h_{ij}\} = g(a_j, b_j, x_i)$ gizli katman çıkış matrisidir. Böylece, çıkış ağırlık vektörü (4.4)'te gösterildiği gibi gizli katman çıkış matrisinin Moore-Penrose genelleştirilmiş tersiyle analitik olarak hesaplanabilir;

$$\hat{\beta} = H^+ T \tag{4.4}$$

burada H^+, H matrisinin Moore-Pensore genelleştirilmiş tersidir.

4.2.2. Ağırlıklandırılmış Aşırı Öğrenme Makinesi

$x_i \in R^n$ ve t_i'nin sınıf etiketleri olduğu, iki sınıfa sahip $[x_i, t_i]$, $i = 1, \dots, N$ bir eğitim veri kümesi düşünelim. İkili sınıflandırmada, t_i ya -1 ya da $+1$ olur. Daha sonra, her biri bir eğitim örneği x_i ile ilişkilendirilen $N \times N$ köşegen matrisi W_{ii} düşünülür. Ağırlıklandırma işlemi genellikle, azınlık sınıfından gelen x_i 'ye daha büyük W_{ii} atanmasıdır. Sınırdaki (marginal) uzaklığı en üst düzeye çıkarmak ve ağırlıklandırılmış toplam hatayı asgari düzeye indirmek için bir optimizasyon problemi kullanılmıştır:

Azalt (Minimize): $\|H\beta - T\|^2$ ve $\|\beta\|$ **(4.5)**

Ayrıca;

Azalt (Minimize): $L_{AÖM} = \frac{1}{2}\|\beta\|^2 + CW\frac{1}{2}\sum_{i=1}^{N}\|\xi_i\|^2$ **(4.6)**

Amaç (Subject to): $h(x_i)\beta = t_i^T - \xi_i^T,\ \ i = 1, \dots, N$ **(4.7)**

burada $T = [t_1, \dots, t_N]$, ξ_i hata vektörü, $h(x_i)$, x_i ve β'ya göre gizli katmandaki özellik eşleme vektörüdür. İkili optimizasyon problemi, Lagrange çarpanı ve Karush-Kuhn-Tucker teoremi kullanılarak çözülebilir. Böylece, gizli katman çıkış ağırlık vektörü β, sol pseudo-invers ya da sağ pseudo-invers ile ilgili olarak (4.7)'den türetilebilir. Küçük boyutlu veriler bulunduğu zaman sağ pseudo-invers tercih edilir çünkü o $N \times N$ matrisinin tersini içermektedir. Aksi halde, sol pseudo-invers daha uygundur çünkü L, N'den çok küçük olduğunda $L \times L$ boyutundaki matrisin tersini hesaplamak daha kolaydır;

N küçük ise: $\beta = H^T\left(\frac{I}{C} + WHH^T\right)^{-1}WT$ **(4.8)**

N büyük ise: $\beta = H^T\left(\frac{I}{C} + H^TWH\right)^{-1}H^TWT$ **(4.9)**

AAÖM'de araştırmacılar iki farklı ağırlıklandırma şeması benimsemiştir. Birincisinde, azınlık ve çoğunluk sınıflarının ağırlıkları şu şekilde hesaplanır;

$$W_{azınlık} = {}^{1}/_{\#(t_i^+)} \text{ ve } W_{cogunluk} = {}^{1}/_{\#(t_i^-)} \quad \textbf{(4.10)}$$

ve ikincisi için, ilgili ağırlıklar aşağıdaki gibi hesaplanır;

$$W_{azınlık} = {}^{0{,}618}/_{\#(t_i^+)} \text{ ve } W_{cogunluk} = {}^{1}/_{\#(t_i^-)} \quad \textbf{(4.11)}$$

4.2.3. NAAÖM: Nötrozofik Ağırlıklandırılmış Aşırı Öğrenme Makinesi

Ağırlıklandırılmış AÖM, azınlık sınıfındaki tüm örneklere aynı ağırlık değerini ve çoğunluk sınıfındaki tüm örneklere de başka bir aynı ağırlık değerini atamaktadır. Bu işlem, bazı dengesiz veri kümelerinde oldukça iyi çalışsa da, bir sınıftaki tüm örneklere aynı ağırlık değerini atamak, gürültüye ve aykırı örneklere sahip olan veri kümeleri için iyi bir seçim olmayabilir. Diğer bir deyişle, dengesiz bir veri kümesindeki gürültülü ve aykırı veri örnekleriyle başa çıkabilmek için, her sınıftaki her bir veriye o verinin sınıfındaki önemini

yansıtan farklı ağırlık değeri vermek gerekmektedir. Bu nedenle, sınıfındaki her bir örneğin önemini belirlemek için yeni bir yöntem sunulmuştur. NCO kümeleme, bir veri örneğinin aitlik, gürültü ve aykırı üyeliklerini belirleyebilmektedir. Daha sonra bu üyelikler, o verinin ağırlık değerinin hesaplanması sırasında kullanılacaktır.

Guo ve Şengür, NCO [13] kümeleme algoritmalarını NK teoremine dayanarak önermişlerdir [66, 67, 108]. NCO yönteminde, BCO yönteminin gürültü ve aykırı veri noktaları üzerindeki zayıflığının üstesinden gelebilmek için yeni bir amaç fonksiyonu geliştirilmiştir. NCO algoritmasında, hem gürültü hem de aykırı değer reddi için iki yeni ret türü geliştirilmiştir. NCO'nun amaç fonksiyonu, "*3.2. Nötrozofik C-Ortalamalar Kümeleme"* bölümünde denklemlerle detaylı olarak tanımlanmıştır. O bölümdeki tanımlar altında, her bir azınlık ve çoğunluk sınıfındaki her giriş örneği bir üçlü T_{ij}, I_i, F_i ile ilişkilendirilmiştir. T_{ij}, örneğin yüksek olasılıkla etiketli sınıfa dahil olduğu anlamına gelirken, I_i, örneğin yüksek olasılıkla belirsiz olduğu anlamına gelmektedir. Son olarak, F_i, örneğin gürültü ya da aykırı veri olduğu anlamına gelmektedir.

NCO'da kümeleme işlemi uygulandıktan sonra, azınlık ve çoğunluk sınıflarının her bir örneği için ağırlıklar aşağıdaki gibi elde edilmektedir:

$$W_{ii}^{azınlık} = {C_r}/{(T_{ij} + I_i - F_i)} \text{ ve } W_{ii}^{cogunluk} = {1}/{(T_{ij} + I_i - F_i)} \quad \textbf{(4.12)}$$

$$C_r = \frac{\#(t_i^-)}{\#(t_i^+)} \quad \textbf{(4.13)}$$

burada C_r, çoğunluk sınıfındaki örnek sayısının azınlık sınıfındaki örnek sayısına oranıdır.

NAAÖM algoritması dört adımdan oluşmaktadır. İlk adım, giriş örneklerinin sınıf etiketlerine göre önceden hesaplanmış küme merkezlerine dayalı NCO algoritmasının uygulanmasını gerektirmektedir. Böylece, bir sonraki adım için T, I ve F üyelik değerleri belirlenmiş olur. Algoritmanın ikinci adımındaki ilgili ağırlıklar belirlenmiş T, I ve F değerlerinden hesaplanmaktadır. Adım 3'te, AÖM parametreleri ayarlanmakta ve H matrisini hesaplamak için örnekler ve ağırlıklar AÖM'ye giriş olarak verilmektedir. Gizli katman ağırlık vektörü β; H, W ve sınıf etiketlerine göre hesaplanmaktadır. Son olarak, test veri kümesinin etiketlerinin belirlenmesi algoritmanın son adımında gerçekleştirilmektedir.

NAAÖM algoritması şu şekildedir:

Giriş:	Etiketlenmiş eğitim veri kümesi
Çıkış:	Tahmin edilen sınıf etiketleri
Adım 1:	Etiketlenmiş veri kümesine göre küme merkezlerini başlat ve her bir veri noktasının T, I ve F değerlerini elde etmek için NCO algoritmasını çalıştır;
Adım 2:	Denklem (4.12) ve (4.13)'e göre $W_{ii}^{azınlık}$ ve $W_{ii}^{cogunluk}$'u hesapla;
Adım 3:	AÖM parametrelerini uyarla ve NAAÖM'yi çalıştır. H matrisini hesapla ve Denklem (4.8) veya Denklem (4.9)'a göre β'yı elde et;
Adım 4:	β'ya dayanarak test veri kümesinin etiketlerini hesapla.

4.3. Deneysel Çalışmalar

Önerilen NAAÖM yönteminin başarımını değerlendirmek için geometrik ortalama (G_{ort}) kullanılmıştır. G_{ort} aşağıdaki şekilde hesaplanmaktadır:

$$G_{ort} = \sqrt{R\frac{DN}{DN + YP}} \tag{4.14}$$

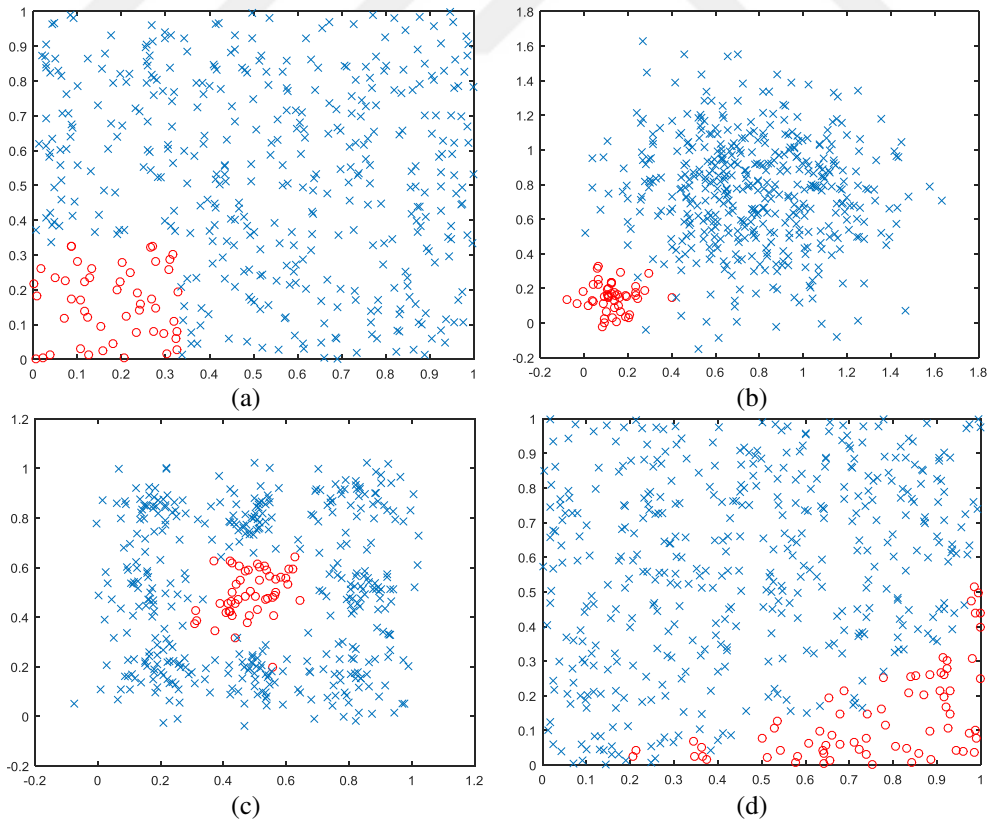
$$R = \frac{DP}{DP + YN} \tag{4.15}$$

burada R, geri çağırma oranını belirtmektedir ve DP (Doğru-Pozitif), DN (Doğru-Negatif), YP (Yanlış-Pozitif), YN 'yi (Yanlış-Negatif) göstermektedir. G_{ort} değeri $[0-1]$ aralığındadır ve pozitif sınıf doğruluk ve negatif sınıf doğruluklarının karekökünü temsil etmektedir. NAAÖM sınıflandırıcısının başarım değerlendirmesi, hem yapay veri kümesinde hem de gerçek veri kümesinde test edilmiştir. Deneylerde beş-katlı çapraz-doğrulama yöntemi benimsenmiştir. NAAÖM'nin gizli düğümünde RTF çekirdeği kabul edilmiştir. Izgara araması, C takas (trade-off) sabiti $\{2^{-18}, 2^{-16}, \dots, 2^{48}, 2^{50}\}$ aralığında ve L gizli düğüm sayısı $\{10, 20, \dots, 1990, 2000\}$ aralığında beş-katlı çapraz-doğrulamayı kullanılarak en iyi sonucu elde etmek için yürütülmüştür. Gerçek veri kümeleri için giriş özniteliklerinin $[-1, 1]$'e normalize edilmiştir. Buna ek olarak, NCO için deneme yanılma yöntemi ile elde edilen $\varepsilon = 10^{-5}$, $\varpi_1 = 0{,}75$, $\varpi_2 = 0{,}125$ ve $\varpi_3 = 0{,}125$ gibi

parametreler seçilmiştir. NCO yönteminin δ parametresi $\{2^{-10}, 2^{-8}, \ldots, 2^{8}, 2^{10}\}$ aralığında araştırılmıştır.

4.3.1. Yapay Veri Kümeleri Üzerindeki Deneyler

Önerilen NAAÖM şemasının sınıflandırma başarımını değerlendirmek için 4 adet 2-sınıflı yapay dengesiz veri kümesi kullanılmıştır. Veri kümesi Şekil 4.1'de gösterilmiştir [109]. Sınıflar arasındaki karar verme sınırı karmaşıktır. Şekil 4.1 (a)'da, düzgün bir dağılım izleyen ilk yapay veri kümesi gösterilmiştir. Görülebileceği gibi, Şekil 4.1 (a)'daki kırmızı renkli daireler azınlık sınıfına aittir ve çoğunluk sınıfı olarak geri kalan veri örnekleri mavi renkli çarpı ile gösterilmiştir. İkinci dengesiz veri kümesi, yani "*Gaussian-1*", Şekil 4.1 (b)'de gösterildiği gibi 1:9 oranındaki örneklerle iki Gauss dağılımı kullanılarak elde edilmiştir. Kırmızı renkli daireler azınlık sınıfını gösterirken, mavi renkli çarpı örnekleri çoğunluk sınıfını göstermektedir. Şekil 4.1 (c)'de, "*Gaussian-2*" gibi başka bir Gauss dağılımına dayanan dengesiz veri kümesi verilmiştir. Bu veri kümesi, aynı sayıda örneğin 3×3 ızgarada düzenlenmiş 9 Gauss dağılımından oluşmuştur. Ortadaki kırmızı renkli daireler azınlık



Şekil 4.1. Dört adet 2B yapay dengesiz veri kümesi (X_1, X_2) (a) Uniform, (b) Gaussian-1, (c) Gaussian-2, (d) Complex

sınıfına aitken, mavi renkli çarpı örnekler çoğunluk sınıfına aittir. Son olarak, Şekil 4.1 (d)'de son yapay dengesiz veri kümesi gösterilmiştir. Azınlık ve çoğunluk sınıfları için örneklerin oranı 1:9 olduğu için karmaşık bir veri kümesi olarak bilinmektedir.

Tablo 4.1 dört veri kümesi üzerinde iki yöntem tarafından 10 adet bağımsız çalıştırmayla ulaşılan G_{ort}'yı göstermektedir. "*Gaussian-1*", "*Gaussian-2*" ve "*Uniform*" yapay veri kümeleri için önerilen NAAÖM yöntemi, ağırlıklandırılmış AÖM şemasına kıyasla daha iyi sonuçlar vermiştir; ancak "*Complex*" yapay veri kümesinde ağırlıklandırılmış AÖM daha iyi sonuç elde etmiştir. Tablodaki daha iyi sonuçlar kalın yazı tipi stili olarak gösterilmiştir. "*Gaussian-2*" veri kümesi için NAAÖM'nin tüm denemelerde daha yüksek bir G_{ort} elde ettiğini belirtmek gerekir.

Tablo 4.1. AAÖM'nin NAAÖM ile yapay veri kümeleri üzerindeki karşılaştırması

Veri Kümesi	AAÖM	NAAÖM	Veri Kümesi	AAÖM	NAAÖM
	G_{ort}	G_{ort}		G_{ort}	G_{ort}
Gaussian-1-1	0,9811	**0,9822**	Gaussian-2-1	0,9629	**0,9734**
Gaussian-1-2	0,9843	**0,9855**	Gaussian-2-2	0,9551	**0,9734**
Gaussian-1-3	0,9944	**0,9955**	Gaussian-2-3	0,9670	**0,9747**
Gaussian-1-4	0,9866	**0,9967**	Gaussian-2-4	0,9494	**0,9649**
Gaussian-1-5	**0,9866**	0,9833	Gaussian-2-5	0,9467	**0,9724**
Gaussian-1-6	**0,9899**	0,9685	Gaussian-2-6	0,9563	**0,9720**
Gaussian-1-7	**0,9833**	0,9685	Gaussian-2-7	0,9512	**0,9629**
Gaussian-1-8	0,9967	**0,9978**	Gaussian-2-8	0,9644	**0,9785**
Gaussian-1-9	**0,9944**	0,9798	Gaussian-2-9	0,9441	**0,9559**
Gaussian-1-10	0,9846	**0,9898**	Gaussian-2-10	0,9402	**0,9623**
Uniform-1	0,9836	**0,9874**	Complex-1	**0,9587**	0,9481
Uniform-2	**0,9798**	0,9750	Complex-2	**0,9529**	0,9466
Uniform-3	0,9760	**0,9823**	Complex-3	0,9587	**0,9608**
Uniform-4	0,9811	**0,9836**	Complex-4	**0,9482**	0,9061
Uniform-5	0,9811	**0,9823**	Complex-5	**0,9587**	0,9297
Uniform-6	0,9772	0,9772	Complex-6	0,9409	**0,9599**
Uniform-7	**0,9734**	0,9403	Complex-7	**0,9644**	0,9563
Uniform-8	0,9785	**0,9812**	Complex-8	**0,9575**	0,9553
Uniform-9	**0,9836**	0,9762	Complex-9	**0,9551**	0,9446
Uniform-10	0,9695	**0,9734**	Complex-10	0,9351	**0,9470**

4.3.2. Gerçek Veri Kümeleri Üzerindeki Deneyler

Bu bölümde, önerilen NAAÖM yönteminin gerçek veri kümeleri üzerindeki başarımı test edilmiştir [110]. Farklı sayıda öznitelik, eğitim örnekleri, test örnekleri ve dengesizlik oranına sahip toplam 21 veri kümesi Tablo 4.2'de gösterilmiştir. Seçilen veri kümeleri, dengesizlik oranlarına göre iki sınıfa ayrılabilir. Birinci sınıf dengesizlik oranı aralığı 0 ile 0,2 arasındadır ve "*yeast-1-2-8-9_vs_7*", "*abalone9_18*", "*glass-0-1-6_vs_2*", "*vowel0*", "*yeast-0-5-6-7-9_vs_4*", "*page-blocks0*", "*yeast3*", "*ecoli2*", "*new-thyroid1*" ve "*new-thyroid2*" veri kümelerini içermektedir. Öte yandan, ikinci sınıf "*ecoli1*", "*glass-0-1-2-3_vs_4-5-6*", "*vehicle0*", "*vehicle1*", "*haberman*", "*yeast*", "*glass0*", "*iris0*", "*pima*", "*wisconsin*", ve "*glass1*" gibi veri kümelerini içermektedir; dengesizlik oranı 0,2 ile 1 arasında değişmektedir.

Tablo 4.2. Gerçek veri kümeleri ve nitelikleri

Veri Kümesi	Öznitelik Sayısı	Eğitim Verisi Sayısı	Test Verisi Sayısı	Dengesizlik Oranı
yeast-1-2-8-9_vs_7	8	757	188	0,0327
abalone9_18	8	584	147	0,0600
glass-0-1-6_vs_2	9	153	39	0,0929
vowel0	13	790	198	0,1002
yeast-0-5-6-7-9_vs_4	8	422	106	0,1047
page-blocks0	10	4377	1095	0,1137
yeast3	8	1187	297	0,1230
ecoli2	7	268	68	0,1806
new-thyroid1	5	172	43	0,1944
new-thyroid2	5	172	43	0.1944
ecoli1	7	268	68	0,2947
glass-0-1-2-3_vs_4-5-6	9	171	43	0,3053
vehicle0	18	676	170	0,3075
vehicle1	18	676	170	0,3439
haberman	3	244	62	0,3556
yeast1	8	1187	297	0,4064
glass0	9	173	43	0,4786
iris0	4	120	30	0,5000
pima	8	614	154	0,5350
wisconsin	9	546	137	0,5380
glass1	9	173	43	0,5405

Yapılan deneysel çalışmalar ile önerilen NAAÖM'nin, AÖM, AAÖM ve DVM ile karşılaştırma sonuçları Tablo 4.3'te verilmiştir. AAÖM yöntemi farklı ağırlıklandırma şeması (W_1, W_2) kullandığından karşılaştırmada daha yüksek G_{ort} değeri üretmiştir. Tablo 4.3'te görüldüğü gibi, NAAÖM yöntemi dengesiz veri kümelerinin 17'sinde daha yüksek G_{ort} değerleri vermiştir. Üç veri kümesi için yüksek G_{ort} değeri paylaşılmıştır. Sadece "*page-blocks0*" veri kümesi için, AAÖM yöntemi daha iyi sonuç vermiştir.

Tablo 4.3. İkili veri kümelerinin G_{ort} açısından deneysel sonucu

	G_{ort}	Gauss çekirdek				Radyal tabanlı çekirdek		
		AÖM		AAÖM encok(W_1,W_2)		DVM	NAAÖM (önerilen yöntem)	
		C	G_{ort} (%)	C	G_{ort} (%)	G_{ort} (%)	C	G_{ort} (%)
dengesizlik oranı: 0 - 0,2	yeast-1-2-8-9_vs_7 (0,0327)	2^{48}	60,97	2^{4}	71,41	47,88	2^{-7}	**77.57**
	abalone9_18 (0,0600)	2^{18}	72,71	2^{28}	89,76	51,50	2^{23}	**94,53**
	glass-0-1-6_vs_2 (0,0929)	2^{50}	63,20	2^{32}	83,59	51,26	2^{7}	**91,86**
	vowel0 (0,1002)	2^{-18}	100,00	2^{-18}	100,00	99,44	2^{7}	100,00
	yeast-0-5-6-7-9_vs_4 (0,1047)	2^{-6}	68,68	2^{4}	82,21	62,32	2^{-10}	**85,29**
	page-blocks0 (0,1137)	2^{4}	89,62	2^{16}	**93,61**	87,72	2^{20}	93,25
	yeast3 (0,1230)	2^{44}	84,13	2^{48}	93,11	84,71	2^{3}	**93,20**
	ecoli2 (0,1806)	2^{-18}	94,31	2^{8}	94,43	92,27	2^{10}	**95,16**
	new-thyroid1 (0,1944)	2^{0}	99,16	2^{14}	99,72	96,75	2^{7}	**100,00**
	new-thyroid2 (0,1944)	2^{2}	99,44	2^{12}	99,72	98,24	2^{7}	**100,00**
dengesizlik oranı: 0,2 - 1	ecoli1 (0,2947)	2^{0}	88,75	2^{10}	91,04	87,73	2^{20}	**92,10**
	glass-0-1-2-3_vs_4-5-6 (0,3053)	2^{10}	93,26	2^{-18}	95,41	91,84	2^{7}	**95,68**
	vehicle0 (0,3075)	2^{8}	99,36	2^{20}	99,36	96,03	2^{10}	99,36
	vehicle1 (0,3439)	2^{18}	80,60	2^{24}	86,74	66,04	2^{10}	**88,06**
	haberman (0,3556)	2^{42}	57,23	2^{14}	66,26	37,35	2^{7}	**67,34**
	yeast1 (0,4064)	2^{0}	65,45	2^{10}	73,17	61,05	2^{10}	**73,19**
	glass0 (0,4786)	2^{0}	85,35	2^{0}	85,65	79,10	2^{13}	**85,92**
	iris0 (0,5000)	2^{-18}	100,00	2^{-18}	100,00	98,97	2^{10}	100,00
	pima (0,5350)	2^{0}	71,16	2^{8}	75,58	70,17	2^{10}	**76,35**
	wisconsin (0,5380)	2^{-2}	97,18	2^{8}	97,70	95,67	2^{7}	**98,22**
	glass1 (0,5405)	2^{-18}	77,48	2^{2}	80,35	69,64	2^{17}	**81,77**

NAAÖM yönteminin, 4 veri kümesi ("*vowel0*", "*new-thyroid1*", "*new-thyroid2*", "*iris0*") için %100 G_{ort} değerine ulaştığını belirtmek gerekir. Buna ek olarak, NAAÖM tüm veri kümeleri için DVM'den daha yüksek G_{ort} değerleri üretmiştir.

Elde edilen sonuçlar, ayrıca eğri altındaki alan (EAA) değerleri ile de değerlendirilmiştir [111]. Buna ek olarak, Tablo 4.4'te gösterildiği üzere önerilen NAAÖM

Tablo 4.4. İkili veri kümelerinin ortalama EAA açısından deneysel sonucu

		Gauss çekirdek				**Radyal tabanlı çekirdek**		
		AÖM		**AAÖM encok(W_1,W_2)**		**DVM**	**NAAÖM (önerilen yöntem)**	
		C	EAA (%)	**C**	EAA (%)	EAA (%)	**C**	EAA (%)
dengesizlik oranı: 0 – 0,2	yeast-1-2-8-9_vs_7 (0,0327)	2^{48}	61,48	2^{4}	65,53	56,67	2^{-7}	**74,48**
	abalone9_18 (0,0600)	2^{18}	73,05	2^{28}	89,28	56,60	2^{23}	**95,25**
	glass-0-1-6_vs_2 (0,0929)	2^{50}	67,50	2^{32}	61,14	53,05	2^{7}	**93,43**
	vowel0 (0,1002)	2^{-18}	93,43	2^{-18}	99,22	99,44	2^{7}	**99,94**
	yeast-0-5-6-7-9_vs_4 (0,1047)	2^{-6}	66,35	2^{4}	80,09	69,88	2^{-10}	**82,11**
	page-blocks0 (0,1137)	2^{4}	67,42	2^{16}	71,55	88,38	2^{20}	**91,49**
	yeast3 (0,1230)	2^{44}	69,28	2^{48}	90,92	83,92	2^{3}	**93,15**
	ecoli2 (0,1806)	2^{-18}	71,15	2^{8}	94,34	92,49	2^{10}	**94,98**
	new-thyroid1 (0,1944)	2^{0}	90,87	2^{14}	98,02	96,87	2^{7}	**100,00**
	new-thyroid2 (0,1944)	2^{2}	84,29	2^{12}	96,63	98,29	2^{7}	**100,00**
dengesizlik oranı: 0,2 - 1	ecoli1 (0,2947)	2^{0}	66,65	2^{10}	90,28	88,16	2^{20}	**92,18**
	glass-0-1-2-3_vs_4-5-6 (0,3053)	2^{10}	88,36	2^{-18}	93,94	92,02	2^{7}	**95,86**
	vehicle0 (0,3075)	2^{8}	71,44	2^{20}	62,41	96,11	2^{10}	**98,69**
	vehicle1 (0,3439)	2^{18}	58,43	2^{24}	51,80	69,10	2^{10}	**88,63**
	haberman (0,3556)	2^{42}	68,11	2^{14}	55,44	54,05	2^{7}	**72,19**
	yeast1 (0,4064)	2^{0}	56,06	2^{10}	70,03	66,01	2^{10}	**73,66**
	glass0 (0,4786)	2^{0}	74,22	2^{0}	75,99	79,81	2^{13}	**81,41**
	iris0 (0,5000)	2^{-18}	100,00	2^{-18}	100,00	99,00	2^{10}	**100,00**
	pima (0,5350)	2^{0}	59,65	2^{8}	50,01	71,81	2^{10}	**75,21**
	wisconsin (0,5380)	2^{-2}	83,87	2^{8}	80,94	95,68	2^{7}	**98,01**
	glass1 (0,5405)	2^{-18}	75,25	2^{2}	80,46	72,32	2^{17}	**81,09**

yöntemi elde edilen EAA değerlerine dayanılarak AÖM, AAÖM ve DVM ile karşılaştırılmıştır. İncelenen tüm veri kümeleri için, önerilen yöntemin EAA değerleri diğer yöntemlere kıyasla daha yüksek başarım sağladığı görülmüştür.

Önerilen NAAÖM yönteminin, AÖM, AAÖM ve DVM yöntemleri ile uygun olarak karşılaştırılması için EAA sonuçlarına ilişkin istatistiksel testler düşünülmüş ve eşleştirilmiş *t*-testi seçilmiştir [112]. Önerilen yöntemin EAA sonuçları ile karşılaştırılan her bir yöntemin EAA sonuçları arasındaki eşleştirilmiş *t*-testi değerleri, *p*-değeri açısından Tablo 4.5'te gösterilmiştir.

Tablo 4.5. Her bir yöntem ile önerilen yöntemin EAA sonuçları arasındaki Eşleştirilmiş *t*-testi değerleri

	Veri Kümesi (dengesizlik oranı)	**AÖM**	**AAÖM**	**DVM**
dengesizlik oranı: 0 – 0,2	yeast-1-2-8-9_vs_7 (0,0327)	**0,0254**	0,0561	**0,0018**
	abalone9_18 (0,0600)	**0,0225**	0,0832	**0,0014**
	glass-0-1-6_vs_2 (0,0929)	**0,0119**	**0,0103**	**0,0006**
	vowel0 (0,1002)	**0,0010**	0,2450	0,4318
	yeast-0-5-6-7-9_vs_4 (0,1047)	**0,0218**	0,5834	0,0568
	page-blocks0 (0,1137)	**0,0000**	**0,0000**	**0,0195**
	yeast3 (0,1230)	**0,0008**	**0,0333**	**0,0001**
	ecoli2 (0,1806)	**0,0006**	0,0839	0,0806
	new-thyroid1 (0,1944)	**0,0326**	0,2089	0,1312
	new-thyroid2 (0,1944)	**0,0029**	0,0962	0,2855
dengesizlik oranı: 0,2 - 1	ecoli1 (0,2947)	**0,0021**	0,1962	0,0744
	glass-0-1-2-3_vs_4-5-6 (0,3053)	**0,0702**	0,4319	**0,0424**
	vehicle0 (0,3075)	**0,0000**	**0,0001**	0,0875
	vehicle1 (0,3439)	**0,0000**	**0,0000**	**0,0001**
	haberman (0,3556)	0,1567	**0,0165**	**0,0007**
	yeast1 (0,4064)	**0,0001**	0,0621	**0,0003**
	glass0 (0,4786)	**0,0127**	0,1688	0,7072
	iris0 (0,5000)	NaN	NaN	0,3739
	pima (0,5350)	**0,0058**	**0,0000**	**0,0320**
	wisconsin (0,5380)	**0,0000**	**0,0002**	**0,0071**
	glass1 (0,5405)	**0,0485**	0,8608	**0,0293**

Tablo 4.5'te önerilen yönteme önemli bir avantaj sağlayan *p*-değerlerinin 0,05'e eşit veya daha küçük olduğu sonuçlar kalın yazı tipi stilinde gösterilmiştir. Böylece önerilen NAAÖM

yöntemiyle, her veri kümesi ve yöntem çiftleri ayrı olarak değerlendirildiğinde toplam 63 testten 39'unda diğer yöntemlerden daha iyi sonuç elde edilmiştir.

EAA değerlerine göre elde edilen sonuçların karşılaştırılması için başka bir istatistiksel test, Friedman hizalı sıralama testi uygulanmıştır [113]. Bu test parametrik olmayan bir testtir ve Holm yöntemi, post-hoc kontrol yöntemi olarak seçilmiştir. Anlamlılık düzeyi 0,05 olarak belirlenmiştir. İstatistikler, STAC [114] aracı ile elde edilmiştir ve Tablo 4.6'da kaydedilmiştir. Bu sonuçlara göre, en yüksek sıra değeri önerilen NAAÖM yöntemi ile elde edilmiştir. DVM ve AAÖM'nin sıra değerleri AÖM'den daha büyüktür. Buna ek olarak, karşılaştırma istatistikleri, ayarlanmış (düzeltilmiş) p-değerleri ve hipotez sonuçları Tablo 4.6'da verilmiştir.

Tablo 4.6. Friedman hizalı sıralama testi (anlamlılık düzeyi 0,05)

İstatistik	**p-değeri**	**Sonuç**
29,6052	0,0000	H0 reddedildi
	Sıralama	
Algoritma		**Sıra**
AÖM		21,7619
AAÖM		38,9047
DVM		41,5238
NAAÖM		67,8095

Karşılaştırma	**İstatistik**	**Ayarlanmış p-değeri**	**Sonuç**
NAAÖM - AÖM	6,1171	0,0000	H0 reddedildi
NAAÖM - AAÖM	3,8398	0,0003	H0 reddedildi
NAAÖM - DVM	3,4919	0,0005	H0 reddedildi

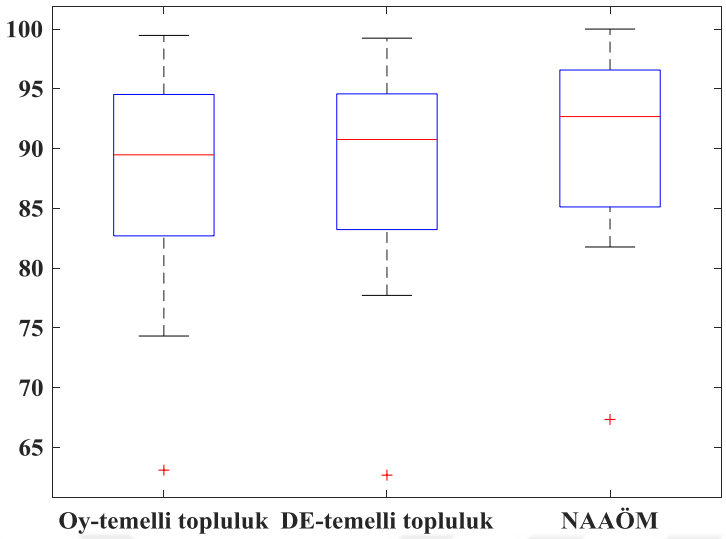
Daha sonra önerilen NAAÖM yöntemi, 12 veri kümesi üzerinde topluluk temelli iki AAÖM yöntemiyle karşılaştırılmıştır [44]. Elde edilen sonuçlar ve ortalama sınıflandırma değerleri G_{ort} Tablo 4.7'de gösterilmiştir. Her bir veri kümesi için en iyi sınıflandırma sonucu kalın yazı tipi stili olarak vurgulanmıştır. Ortalama sınıflandırma başarımı üzerine genel bir bakış yapılırsa, NAAÖM yönteminin her iki topluluk temelli AAÖM yöntemlerine göre en yüksek G_{ort} değerini verdiği görülmektedir. Buna ek olarak, önerilen NAAÖM yöntemi, "*yeast3*" ve "*glass2*" veri kümeleri istisna olarak, 12 veri kümesinin 10'unda G_{ort} değerlerine göre diğer iki yöntemden daha üstündür.

Dikkatli bir gözlem ile görülebileceği üzere, NAAÖM yöntemi "*glass1*", "*haberman*", "*yeast1_7*" ve "*abalone9_18*" veri kümeleri açısından başarımı önemli ölçüde geliştirmemiştir ancak her iki topluluk temelli AAÖM yönteminden biraz daha iyi başarım göstermiştir.

Tablo 4.7. Önerilen yöntemin topluluk temelli iki AAÖM yöntemiyle karşılaştırılması

Veri Kümesi	Oy-temelli topluluk		DE-temelli topluluk		NAAÖM	
	C	G_{ort} (%)	C	G_{ort} (%)	C	G_{ort} (%)
glass1	2^{30}	74,32	2^{18}	77,72	2^{17}	**81,77**
haberman	2^{12}	63,10	2^{28}	62,68	2^{7}	**67,34**
ecoli1	2^{40}	89,72	2^{0}	91,39	2^{20}	**92,10**
new-thyroid2	2^{10}	99,47	2^{32}	99,24	2^{7}	**100,00**
yeast3	2^{4}	94,25	2^{2}	**94,57**	2^{3}	93,20
ecoli3	2^{10}	88,68	2^{18}	89,50	2^{17}	**92,16**
glass2	2^{8}	86,45	2^{16}	**87,51**	2^{7}	85,58
yeast1_7	2^{20}	78,95	2^{38}	78,94	2^{-6}	**84,66**
ecoli4	2^{8}	96,33	2^{14}	96,77	2^{10}	**98,85**
abalone9_18	2^{4}	89,24	2^{16}	90,13	2^{23}	**94,53**
glass5	2^{18}	94,55	2^{12}	94,55	2^{7}	**95,02**
yeast5	2^{12}	94,51	2^{28}	94,59	2^{17}	**98,13**
Ortalama		87,46		88,13		**90,53**

Karşılaştırılan yöntemlerin kutu çizimleri Şekil 4.2'de gösterilmiştir. NAAÖM yöntemi tarafından üretilen kutu, karşılaştırılan oy-temelli topluluk ve diferansiyel evrim (DE)-temelli topluluk yöntemleri tarafından oluşturulan kutulardan daha kısadır. NAAÖM yönteminin dağılım derecesi nispeten daha düşüktür. Ayrıca, tüm yöntemlere ait kutu çizimlerinin "*haberman*" veri kümesinin G_{ort} 'unu bir istisna olarak değerlendirdiğini belirtmek gerekir. Son olarak kutu çizimleri, önerilen NAAÖM yönteminin topluluk temelli AAÖM yöntemlerine kıyasla daha dayanıklı olduğunu belirlemiştir.



Şekil 4.2. Karşılaştırılan yöntemlerin kutu çizimi gösterimi

5. NÖTROZOFİK KÜME TABANLI *k*-EN YAKIN KOMŞULUK SINIFLANDIRICISI

5.1. Giriş

Bu bölümde, NK teorisine dayanan yeni bir basit *k*-EYK yaklaşımı önerilmiştir. Önerilen yöntem Nötrozofik Küme tabanlı *k*-En Yakın Komşuluk (NK-*k*-EYK) olarak adlandırılmıştır. *k*-EYK sınıflandırıcısının sınıflandırma başarımını artırmak için NK üyelikleri benimsenmiştir. Bu amaçla, NCO algoritması, kümelerin merkezlerini elde etmek için etiketlenmiş eğitim verilerinin kullanıldığı eğiticili bir şekilde ele alınmıştır. Nihai aitlik üyeliği derecesi U, NK üçlüsünden $U = T + I - F$ olarak hesaplanmıştır. Bulanık *k*-EYK'da verildiği gibi benzer bir nihai oylama şeması sınıf etiketlerinin belirlenmesi için kullanılmıştır.

Önerilen NK-*k*-EYK yönteminin başarımı kapsamlı deneyler yapılarak değerlendirilmiştir. Bu amaçla birkaç yapay ve gerçek dünya veri kümeleri kullanılmıştır. Önerilen yöntem, *k*-EYK, bulanık *k*-EYK ve iki ağırlıklandırılmış *k*-EYK şemaları ile kıyaslanmıştır. Sonuçlar teşvik edici ve ilerlemeye açıktır.

5.2. Önerilen Yöntem

5.2.1. *k*-En Yakın Komşuluk Sınıflandırıcısı

Literatür bölümünde de belirtildiği gibi, *k*-EYK, 1951 yılında önerilen en basit, popüler, eğiticili ve parametrik olmayan bir sınıflandırma yöntemidir [46]. Test verilerinin eğitim depolarında saklanan veri örneklerine olan benzerliğini ölçen, uzaklığa dayalı bir sınıflandırıcıdır. Daha sonra, test verileri eğitim kümesindeki en yakın k komşusunun çoğunluk oylarıyla etiketlenmektedir.

$X = \{x_1, x_2, \dots, x_N\}$, $x_i \in R^n$'nin n boyutlu öznitelik uzayında eğitim noktalarının bulunduğu eğitim kümesi olduğunu düşünelim ve $Y = \{y_1, y_2, \dots, y_N\}$ bunlara karşılık gelen sınıf etiketleri olsun. Sınıf etiketi bilinmeyen bir test veri noktası $\acute{x}$ verildiğinde, bu aşağıdaki gibi belirlenebilir:

- Bir uzaklık fonksiyonu (Öklid uzaklığı gibi) kullanarak test örneği ile eğitim örneği arasındaki benzerlik ölçümlerini hesaplayınız.

- Benzerlik ölçüsüne göre, test veri örneğinin *k* adet en yakın komşularını eğitim veri örneklerinde bulunuz.

5.2.2. Bulanık *k*-En Yakın Komşuluk Sınıflandırıcısı

k-EYK'da, bir eğitim veri örneği *x*'in verilen sınıflardan birine ait olduğu kabul edildiğinde, o eğitim örneğinin her bir C sınıfına olan U üyeliği $\{0, 1\}$'deki bir değerler diziyle verilir. Eğitim veri örneği x, c_1 sınıfına ait ise, $C = \{c_1, c_2\}$ olduğundan $U_{c_1}(x) = 1$ ve $U_{c_2}(x) = 0$ olur.

Bununla birlikte, bulanık *k*-EYK'da net üyelikler kullanmak yerine bulanık teorinin doğası gereği sürekli aralıkta üyelikler kullanılmaktadır [53]. Böylece eğitim veri örneğinin üyeliği şu şekilde hesaplanabilir:

$$U_{c_1}(x) = \begin{cases} 0.51 + 0.49\dfrac{k_{c_1}}{K} \; eğer \; c = c_1 \; ise \\ 0.49\dfrac{k_{c_1}}{K} \quad diğer \; durumlarda \end{cases} \tag{5.1}$$

burada k_{c_1}, $\acute{x}$'in *k* komşuları arasında bulunan c_1 sınıfına ait örnek sayısını göstermekte ve *K* ise bir tek tam sayıdır.

Bulanık üyelikler hesaplandıktan sonra bir test örneğinin sınıf etiketi şu şekilde belirlenebilir:

- Test örneğinin *k* en yakın komşularını Öklid uzaklığı ile belirleyiniz.
- Öklid normunu ve üyelikleri kullanarak her bir sınıf ve komşu için bir nihai oy üretiniz.

$$V(k_j, c) = \frac{U_c(k_j) \Big/ (\|\acute{x} - k_j\|)^{2/m-1}}{\sum_{i=1}^{k} 1 \Big/ (\|\acute{x} - k_i\|)^{2/m-1}} \tag{5.2}$$

burada k_j , j 'ninci en yakın komşu ve $m = 2$ bir parametredir. Daha sonra nihai sınıflandırmayı elde etmek için her komşunun oyları eklenir.

5.2.3. NK-*k*-EYK: Nötrozofik Küme tabanlı *k*-En Yakın Komşuluk Sınıflandırıcı

Geleneksel *k*-EYK, eğitim veri kümelerindeki sınıf etiketlerine eşit ağırlıklar atamasından muzdarip olduğundan, bu kısıtlamanın üstesinden gelmek için nötrozofik üyelikler benimsenmiştir. Nötrozofik üyelikler, veri noktasının kendi sınıfındaki önemini yansıtmaktadır ve bu üyelikler *k*-EYK yaklaşımı için yeni bir işlem olarak kullanılabilir.

Nötrozofik küme, bir örneğin doğruya, yanlışa ve belirsizliğe ait üyeliklerini belirleyebilir. Eğiticisiz nötrozofik kümeleme algoritması NCO, eğiticili bir şekilde kullanılabilir [13, 115]. Net (keskin, crisp) kümeleme yöntemleri, her bir veri noktasının kümelerin merkezine olan yakınlığına göre bir kümeye ait olması gerektiğini varsaymaktadır. Bulanık kümeleme yöntemleri, her veri noktasına kümenin merkezine yakınlığına göre bulanık üyelikler atamıştır. Nötrozofik kümeleme, her veri noktasına T, I ve F üyeliklerini sadece bir küme merkezine olan yakınlığına göre değil, aynı zamanda iki kümenin merkez ortalamasına yakınlığına göre atamıştır. NCO kümeleme hakkında ayrıntılı bilgi Bölüm 3.2'de ve kaynak [13]'de bulunmaktadır. Bir eğiticili öğrenmede eğitim veri kümelerinin etiketleri bilindiğinden, kümelerin merkezleri buna göre hesaplanabilir. Sonrasında, doğru (T), belirsizlik (I) ve yanlış (F) üyelikleri aşağıdaki gibi hesaplanabilir:

$$T_{ij} = \frac{(x_i - c_j)^{-\left(\frac{2}{m-1}\right)}}{\sum_{j=1}^{C}(x_i - c_j)^{-\left(\frac{2}{m-1}\right)} + (x_i - \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)} + \delta^{-\left(\frac{2}{m-1}\right)}} \tag{5.3}$$

$$I_i = \frac{(x_i - \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)}}{\sum_{j=1}^{C}(x_i - c_j)^{-\left(\frac{2}{m-1}\right)} + (x_i - \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)} + \delta^{-\left(\frac{2}{m-1}\right)}} \tag{5.4}$$

$$F_i = \frac{(\delta)^{-\left(\frac{2}{m-1}\right)}}{\sum_{j=1}^{C}(x_i - c_j)^{-\left(\frac{2}{m-1}\right)} + (x_i - \bar{c}_{imax})^{-\left(\frac{2}{m-1}\right)} + \delta^{-\left(\frac{2}{m-1}\right)}} \tag{5.5}$$

burada m bir sabit, δ bir düzenleyici parametredir ve c_j , j kümesinin merkezini göstermektedir. Her bir i noktası için, $\bar{c}_{imax}$ doğru üyelik değerlerinin diğerlerinden daha büyük olduğu iki küme merkezinin ortalamasıdır. T_{ij}, j sınıfı için i noktasının doğru üyelik

değerini gösterir. F_i, i noktasının yanlış üyelik değerini göstermekte ve I_i, i noktasının belirsizlik üyelik değerini belirtmektedir. Daha büyük T_{ij} değeri, i noktasının bir kümenin yakınında olduğu ve bu noktanın gürültü olma olasılığının daha az olduğu anlamına gelir. Daha büyük I_i değeri, i noktasının herhangi iki küme arasında olduğunu ve daha büyük F_i değerinin, i noktasının muhtemelen gürültü olduğunu gösterdiğini ifade etmektedir. i noktası için nihai bir üyelik değeri, doğru üyelik değerine belirsizlik üyelik değerini ekleyerek ve (5.6)'da gösterildiği gibi yanlış üyelik değerini çıkararak hesaplanabilir.

Nötrozofik üyelik üçlemesini belirledikten sonra, j. sınıf etiketi için bilinmeyen bir x_u örneğinin üyeliği [54]'te olduğu gibi hesaplanabilir.

$$\mu_{ju} = \frac{\sum_{i=1}^{k} d_i(T_{ij} + I_i - F_i)}{\sum_{i=1}^{k} d_i} \tag{5.6}$$

$$d_i = \frac{1}{\|x_u - x_i\|^{\frac{2}{q-1}}} \tag{5.7}$$

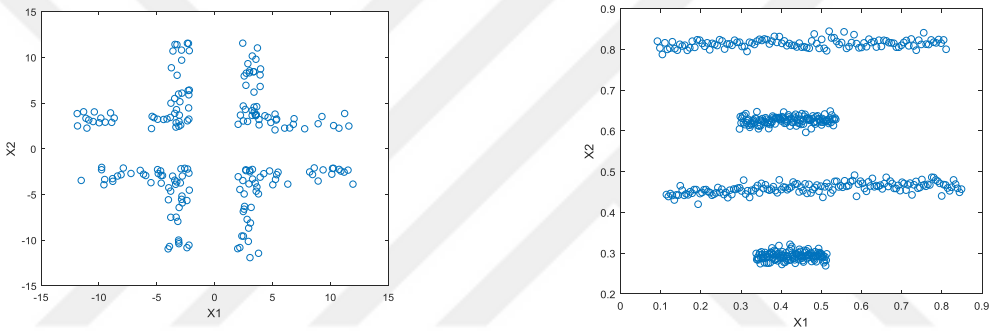
burada d_i, x_i ve x_u arasındaki uzaklığı ölçmeye yarayan uzaklık fonksiyonu, k en yakın komşuluğun değeri, ve q bir tam sayıdır. Bilinmeyen bir x_u örneğinin nötrozofik üyelik değerleri tüm sınıflara atandıktan sonra, NK-*k*-EYK, x_u örneğini nötrozofik üyeliği en fazla olan sınıfa atar. Önerilen NK-*k*-EYK yöntemini oluşturmak için aşağıdaki adımlar uygulanır:

Adım 1:	Etiketlenmiş veri kümesine göre küme merkezlerini başlat ve her bir eğitim veri kümesindeki veri noktasının T, I ve F değerlerini Denklem (5.3), (5.4) ve (5.5)'i kullanarak elde et;
Adım 2:	Test veri örneklerinin üyelik derecelerini Denklem (5.6) ve (5.7)'e göre hesapla;
Adım 3:	Nötrozofik üyeliği en büyük olan sınıfa bilinmeyen test veri noktalarının sınıflarını atayın.

5.3. Deneysel Çalışmalar

Önerilen yöntemi test etmek için iki yapay veri kümesi kullanılmış ve parametrelerin sınıflandırma doğruluğuna olan etkisi incelenmiştir. Öte yandan önerilen yöntem ile geleneksel *k*-EYK ve bulanık *k*-EYK yöntemlerini karşılaştırmak için birkaç gerçek veri kümesi kullanılmıştır. Ayrıca önerilen yöntem, A*k*-EYK ve ÇA*k*-EYK gibi ağırlıklandırılmış *k*-EYK yöntemiyle de karşılaştırılmıştır.

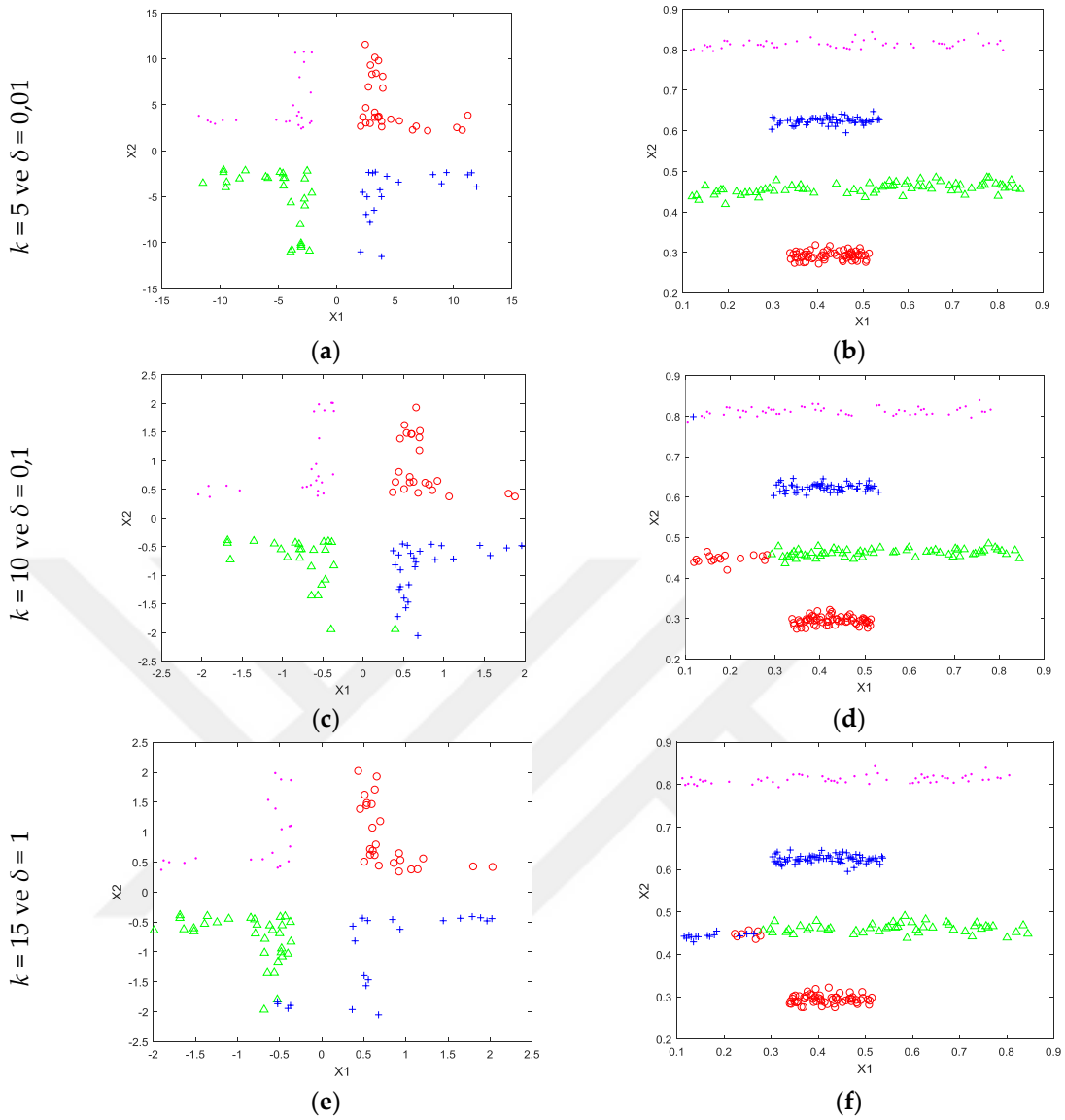
Deneylerde kullanılan "*Corner*" ve "*Line*" yapay veri kümeleri sırasıyla Şekil 5.1 (a) ve (b)'de gösterilmiştir. Her iki yapay veri kümesi dört sınıflı 2B'li veri içermektedir. Yapay veri kümelerinin rastgele seçilen yarısı eğitim için, diğer yarısı test için kullanılmıştır.



(**a**) (**b**)

Şekil 5.1. Yapay veri kümeleri (a) "*Corner*" veri kümesi, (b) "*Line*" veri kümesi

k değeri sırasıyla 5, 10 ve 15; δ değeri sırasıyla 0,01, 0,1 ve 1 olarak seçilmiştir. Elde edilen sonuçlar Şekil 5.2'de sırayla gösterilmiştir. Şekil 5.2'nin ilk satırında görüldüğü gibi, önerilen NK-*k*-EYK yöntemi her iki veri kümesi için de $k = 5$ ve $\delta = 0{,}01$ değerleri ile %100 sınıflandırma doğruluğu elde etmiştir. Bununla birlikte, Şekil 5.2'nin ikinci ve üçüncü satırlarında gösterildiği gibi diğer parametre değerleri için %100 sınıflandırma doğruluğu elde edilememiştir. Bu durum önerilen NK-*k*-EYK yönteminin *k* ve δ uzayında bir parametre ayar mekanizmasına ihtiyacı olduğunu göstermektedir. Bu nedenle k, 2 ile 15 arasında bir tamsayı değerine ayarlanmış ve δ parametresi $\{2^{-10}, 2^{-8}, \ldots, 2^{8}, 2^{10}\}$ aralığında aranmıştır.



Şekil 5.2. "*Corner*" ve "*Line*" veri kümeleri için çeşitli *k* ve δ değerleriyle sınıflandırma sonuçları

KEEL veri kümeleri deposundan [110] indirilen 39 gerçek veri kümesi üzerinde daha ileri deneyler yapılmıştır. Her bir veri kümesi, çapraz doğrulama işlemine (5-katlı veya 10-katlı) göre daha önceden bölümlere ayrılmıştır. Tablo 5.1, her bir veri kümesinin örnek sayısı, öznitelik sayısı ve sınıf sayısı gibi çeşitli özelliklerini göstermektedir. Tüm özellik değerleri $[-1, 1]$'e normalize edilmiş ve tüm deneylerde 5-katlı çapraz doğrulama işlemi benimsenmiştir. Doğruluk değerleri, doğru sınıflandırılmış örnek sayısının toplam örnek sayısına oranı olarak hesaplanmıştır.

Tablo 5.1. KEEL veri deposundan veri kümeleri ve özellikleri [110]

Veri Kümesi	Örnek Sayısı	Öznitelik Sayısı	Sınıf Sayısı	Veri Kümesi	Örnek Sayısı	Öznitelik Sayısı	Sınıf Sayısı
Appendicitis	106	7	2	Penbased	10992	16	10
Balance	625	4	3	Phoneme	5404	5	2
Banana	5300	2	2	Pima	768	8	2
Bands	365	19	2	Ring	7400	20	2
Bupa	345	6	2	Satimage	6435	36	7
Cleveland	297	13	5	Segment	2310	19	7
Dermatology	358	34	6	Sonar	208	60	2
Ecoli	336	7	8	Spectfheart	267	44	2
Glass	214	9	7	Tae	151	5	3
Haberman	306	3	2	Texture	5500	40	11
Hayes-roth	160	4	3	Thyroid	7200	21	3
Heart	270	13	2	Twonorm	7400	20	2
Hepatitis	80	19	2	Vehicle	846	18	4
Ionosphere	351	33	2	Vowel	990	13	11
Iris	150	4	3	Wdbc	569	30	2
Mammographic	830	5	2	Wine	178	13	3
Monk-2	432	6	2	Winequality-red	1599	11	11
Movement	360	90	15	Winequality-white	4898	11	11
New thyroid	215	5	3	Yeast	1484	8	10
Page-blocks	5472	10	5	-	-	-	-

Bulgulara ek olarak, deney sonuçları aynı gerçek veri kümeleri üzerinde yapılan *k*-EYK ve bulanık *k*-EYK deney sonuçları ile de karşılaştırılmıştır. Elde edilen sonuçlar Tablo 5.2'de gösterilmiştir ve en iyi sonuçlar kalın yazı tipi stili olarak belirtilmiştir. Tablo 5.2'de görüldüğü gibi, önerilen NK-*k*-EYK yöntemi 39 veri kümesinden 27'sinde diğer yöntemlerden daha iyi sonuç vermiştir. Buna ek olarak, *k*-EYK ve bulanık *k*-EYK sırasıyla 39 veri kümesindeki 6 ve 7 veri kümesinde daha iyi başarım göstermiştir. Önerdiğimiz yöntem iki veri kümesinde ("*New thyroid*", "*Wine*") %100 başarı elde etmiştir. Üstelik önerilen NK-*k*-EYK yöntemi 13 veri kümesi için %90'nın üzerinde doğruluk değerine ulaşmıştır. Diğer taraftan, "*Winequality-white*" veri kümesi için %33,33'lük daha kötü bir başarım kaydedilmiştir. Ayrıca doğruluk değeri %50'nin altında olan toplam 3 tane veri kümesi mevcuttur.

Tablo 5.2. *k*-EYK ve bulanık *k*-EYK'ya karşı önerilen yöntem NK-*k*-EYK'nın deney sonuçları

Veri Kümesi	*k*-EYK	Bulanık *k*-EYK	Ö. Yöntem NK-*k*-EYK	Veri Kümesi	*k*-EYK	Bulanık *k*-EYK	Ö. Yöntem NK-*k*-EYK
Appendicitis	87,91	**97,91**	90,00	Penbased	99,32	**99,34**	86,90
Balance	89,44	88,96	**93,55**	Phoneme	88,49	**89,64**	79,44
Banana	**89,89**	89,42	60,57	Pima	73,19	73,45	**81,58**
Bands	71,46	70,99	**75,00**	Ring	71,82	63,07	**72,03**
Bupa	62,53	66,06	**70,59**	Satimage	90,94	90,61	**92,53**
Cleveland	56,92	56,95	**72,41**	Segment	95,41	96,36	**97,40**
Dermatology	96,90	96,62	**97,14**	Sonar	83,10	83,55	**85,00**
Ecoli	82,45	83,34	**84,85**	Spectfheart	77,58	78,69	**80,77**
Glass	70,11	72,83	**76,19**	Tae	45,79	67,67	**86,67**
Haberman	71,55	68,97	**80,00**	Texture	**98,75**	**98,75**	80,73
Hayes-roth	30,00	65,63	**68,75**	Thyroid	**94,00**	93,92	74,86
Heart	80,74	80,74	**88,89**	Twonorm	97,11	97,14	**98,11**
Hepatitis	**89,19**	85,08	87,50	Vehicle	**72,34**	71,40	54,76
Ionosphere	96,00	96,00	**97,14**	Vowel	97,78	**98,38**	49,49
Iris	85,18	84,61	**93,33**	Wdbc	97,18	97,01	**98,21**
Mammographic	81,71	80,37	**86,75**	Wine	96,63	97,19	**100,00**
Monk-2	96,29	89,69	**97,67**	Winequality-red	55,60	**68,10**	46,84
Movement	**78,61**	36,11	50,00	Winequality-white	51,04	**68,27**	33,33
New thyroid	95,37	96,32	**100,00**	Yeast	57,62	59,98	**60,81**
Page-blocks	95,91	95,96	**96,34**	-	-	-	-

Daha sonra UCI veri deposundan [116] birkaç veri kümesi üzerinde deneyler gerçekleştirildi. Toplamda, bu deneylerde 11 veri kümesi düşünülmüştür ve sonuçlar iki ağırlıklandırılmış *k*-EYK yaklaşımı olan A*k*-EYK ve ÇA*k*-EYK ile karşılaştırılmıştır. Deneylerde kullanılan UCI veri deposundaki her bir veri kümesinin özellikleri Tablo 5.3'te gösterilmiş ve elde edilen tüm sonuçlar en yüksek doğruluk değerleri kalın yazı tipi sitilinde yazılarak Tablo 5.4'te verilmiştir.

Tablo 5.3. UCI veri deposundan bazı veri kümeleri ve özellikleri [116]

Veri Kümesi	Öznitelik Sayısı	Örnek Sayısı	Sınıf Sayısı	Eğitim Örneği Sayısı	Test Örneği Sayısı
Glass	10	214	7	140	74
Wine	13	178	3	100	78
Sonar	60	208	2	120	88
Parkinson	22	195	2	120	75
Iono	34	351	2	200	151
Musk	166	476	2	276	200
Vehicle	18	846	4	500	346
Image	19	2310	7	1310	1000
Cardio	21	2126	10	1126	1000
Landsat	36	6435	7	3435	3000
Letter	16	20000	26	10000	10000

Tablo 5.4'te görüldüğü gibi, önerilen NK-*k*-EYK yöntemi 11 veri kümesinden 8'inde diğer yöntemlerden daha iyi başarım göstermiştir ve geri kalan 3 veri kümesinde ÇA*k*-EYK daha iyi sonuç vermiştir. "*Parkinson*", "*Image*" ve "*Landsat*" veri kümeleri için, önerilen yöntem %90'nın üzerinde doğruluk değerleri üretmiştir ve en kötü doğruluk sonucu %60,81 ile "*Glass*" veri kümesine aittir. A*k*-EYK ve ÇA*k*-EYK neredeyse aynı doğruluk değerlerini üretmiş ve "*Letter*" ile "*Glass*" veri kümelerinde önerilen yöntemden önemli ölçüde daha iyi başarım göstermiştir.

Tablo 5.4. A*k*-EYK ve ÇA*k*-EYK'ya karşı önerilen yöntem NK-*k*-EYK'nın doğruluk değerleri

Veri Kümesi	A*k*-EYK (%)	ÇA*k*-EYK (%)	Önerilen Yöntem NK-*k*-EYK (%)
Glass	69,86	**70,14**	60,81
Wine	71,47	71,99	**79,49**
Sonar	81,59	82,05	**85,23**
Parkinson	83,53	83,93	**90,67**
Iono	84,27	84,44	**85,14**
Musk	84,77	85,10	**86,50**
Vehicle	63,96	64,34	**71,43**
Image	95,19	95,21	**95,60**
Cardio	70,12	**70,30**	66,90
Landsat	90,63	90,65	**91,67**
Letter	94,89	**94,93**	63,50

Ek olarak, her bir KEEL veri kümesi üzerinde her yöntemin çalışma sürelerini karşılaştırılmıştır. Elde edilen çalışma süreleri Tablo 5.5'te verilmiştir. Deneyler, Intel Core i7-4810 işlemcili ve 32 GB belleğe sahip bir bilgisayarda MATLAB 2014b [117] program ortamında gerçekleştirilmiştir. Tablo 5.5'te görüldüğü gibi, bazı veri kümeleri için k-EYK ve bulanık k-EYK yöntemleri, önerilen NK-*k*-EYK yönteminin ulaştığından daha düşük çalışma sürelerine ulaşmıştır. Ancak ortalama çalışma süreleri göz önüne alındığında, önerilen yöntem en düşük çalışma süresine 0,69 saniyeyle ulaşmıştır. *k*-EYK yöntemi de 1,41 saniye ile ikinci en düşük çalışma süresine sahiptir. Bulanık k-EYK yöntemi, diğer yöntemlerle karşılaştırıldığında ortalama en yavaş çalışma süresini elde etmiştir. Bulanık k-EYK yönteminin çalışma süresi 3,17 saniyede kalmıştır.

Tablo 5.5. Her bir yöntem için çalışma sürelerinin karşılaştırılması

Veri Kümesi	*k*-EYK	Bulanık *k*-EYK	Ö. Yöntem NK-*k*-EYK	Veri Kümesi	*k*-EYK	Bulanık *k*-EYK	Ö. Yöntem NK-*k*-EYK
Appendicitis	0,11	0,16	0,15	Penbased	10,21	18,20	3,58
Balance	0,15	0,19	0,18	Phoneme	0,95	1,88	0,71
Banana	1,03	1,42	0,57	Pima	0,45	0,58	0,20
Bands	0,42	0,47	0,19	Ring	6,18	10,30	2,55
Bupa	0,14	0,28	0,16	Satimage	8,29	15,25	1,96
Cleveland	0,14	0,18	0,19	Segment	1,09	1,76	0,63
Dermatology	0,33	0,31	0,22	Sonar	0,15	0,21	0,23
Ecoli	0,12	0,26	0,17	Spectfheart	0,14	0,25	0,22
Glass	0,10	0,18	0,18	Tae	0,13	0,12	0,16
Haberman	0,13	0,24	0,16	Texture	6,72	12,78	4,30
Hayes-roth	0,07	0,11	0,16	Thyroid	5,86	9,71	2,14
Heart	0,22	0,33	0,17	Twonorm	5,89	10,27	2,69
Hepatitis	0,06	0,06	0,16	Vehicle	0,17	0,31	0,27
Ionosphere	0,13	030	0,25	Vowel	0,47	0,62	0,31
Iris	0,23	0,13	0,16	Wdbc	0,39	0,46	0,26
Mammographic	0,21	0,22	0,20	Wine	0,08	0,14	0,17
Monk-2	0,27	0,33	0,17	Winequality-red	0,28	0,46	0,34
Movement	0,16	0,34	0,35	Winequality-white	1,38	1,95	0,91
New thyroid	0,14	0,18	0,17	Yeast	0,44	0,78	0,30
Page-blocks	1,75	2,20	0,93	Ortalama	1,41	3,17	0,69

Genel olarak, Tablo 5.2'de ve Tablo 5.4'te gösterilen doğruluk değerleri dikkate alındığında önerilen NK-*k*-EYK yönteminin başarılı olduğu söylenebilir. NK-*k*-EYK yönteminin bu yüksek doğruluk değerlerini elde etmesinin sebebi; etkili eğiticili bir sınıflandırıcı oluşturmak için NK teorisi ile uzaklık öğrenmesinin birleştirmesidir. Çalışma zamanı değerlendirmesi de, NK-*k*-EYK'nın diğer ilgili sınıflandırıcılara kıyasla oldukça etkili bir sınıflandırıcı olduğunu kanıtlamıştır.

6. NÖTROZOFİK ÇİZGE KESİM KULLANAN ETKİLİ BİR GÖRÜNTÜ BÖLÜTLEME ALGORİTMASI

6.1. Giriş

Bu bölümde, nötrozofinin görüntü üzerindeki belirsizliği yorumlama avantajı kullanılarak görüntü bölütleme için NK ve çizge kesim (ÇK) birleştirilmiştir. Bunun sonucunda etkili görüntü bölütlemesi için Nötrozofik Çizge Kesim (NÇK) adında yeni bir yöntem önerilmiştir. İlk olarak, görüntü NK kullanılarak yorumlanmış ve buna göre belirsizlik dereceleri hesaplanmıştır. Daha sonra uzamsal ve yoğunluk bilgisinin birleştirilmesiyle tanımlanan görüntü üzerindeki belirsizlik değerini kullanan bir belirsizlik filtresi oluşturulmuştur. Belirsizlik filtresi, uzamsal ve yoğunluk bilgisinin belirsizliğini azaltmaktadır. Görüntü üzerinde bir çizge tanımlanmış ve her bir pikselin ağırlığı, belirsizlik filtresi kullanımından sonraki değer ile gösterilmiştir. Enerji fonksiyonu da nötrozofik değer kullanılarak yeniden tanımlanmıştır. Nihai bölütleme sonuçlarını elde etmek için, çizge üzerinde en büyük akış (maximum-flow) algoritması uygulanmıştır.

Önerilen NÇK yöntemi, hem gürültüsüz hem de farklı gürültü seviyelerine sahip doğal görüntülerde bölütleme uygulaması yapılarak değerlendirilmiştir. Mevcut yöntemlerle karşılaştırıldığında önerilen yöntemin hem nicelik bakımından hem de nitelik bakımından daha iyi başarım gösterdiği ve etkili olduğu görülmüştür.

6.2. Önerilen Yöntem

6.2.1. Nötrozofik Görüntü

NK'deki bir öğe (eleman) şu şekilde tanımlanır: $A = \{A_1, A_2, \dots, A_m\}$ nötrozofik kümede bir alternatif küme olsun. Alternatif A_i, $\{T(A_i), I(A_i), F(A_i)\}/A_i$'dir, burada $T(A_i)$, $I(A_i)$ ve $F(A_i)$ doğruluk, belirsizlik ve yanlışlık kümelerinin üyelik değerleridir.

NK'deki bir I_m görüntüsüne nötrozofik görüntü denmektedir. T_S, I_S ve F_S kullanılarak yorumlanan I_{NS} şeklinde gösterilmektedir. I_{NS}'de bir $P(x,y)$ pikseli verildiğinde $P_{NS}(x,y) = \{T_S(x,y),\ I_S(x,y),\ F_S(x,y)\}$ olarak yorumlanır. $T_S(x,y)$, $I_S(x,y)$ ve $F_S(x,y)$ sırasıyla ön plana, belirsiz kümeye ve arka plana ait üyelikleri temsil etmektedir.

Yoğunluk değeri ve yerel uzamsal bilgiye dayalı olarak, doğruluk ve belirsizlik üyelikleri yerel komşuluklar boyunca belirsizliği tanımlamak için şu şekilde kullanılmıştır:

$$T_S(x,y) = \frac{g(x,y) - g_{min}}{g_{max} - g_{min}} \tag{6.1}$$

$$I_S(x,y) = \frac{G_d(x,y) - G_{d_{min}}}{G_{d_{max}} - G_{d_{min}}} \tag{6.2}$$

burada $g(x,y)$ ve $G_d(x,y)$, görüntü üzerindeki (x,y) pikselinin yoğunluk ve eğim büyüklükleridir.

Ayrıca farklı gruplar arasındaki yoğunluğun belirsizliğini göz önüne alan, küresel yoğunluk dağılımına dayalı nötrozofik üyelik değerleri de hesaplanmıştır. NCO kümeleme diğer algoritmalardaki belirsiz noktaların ele alınmasındaki dezavantajların üstesinden gelmektedir [13]. Burada, bölütlenecek yoğunluktaki farklı gruplar arasında bulunan belirsizlik değerlerini elde etmek için NCO kullanılmıştır.

NCO kullanarak, doğru ve belirsiz üyelikler şu şekilde tanımlanmaktadır:

$$K = \left[\frac{1}{\varpi_1} \sum_{j=1}^{C} (x_i - c_j)^{-\frac{2}{m-1}} + \frac{1}{\varpi_2} (x_i - \bar{c}_{imax})^{-\frac{2}{m-1}} + \frac{1}{\varpi_3} \delta^{-\frac{2}{m-1}} \right]^{-1} \tag{6.3}$$

$$T_{n_{ij}} = \frac{K}{\varpi_1} (x_i - c_j)^{-\frac{2}{m-1}} \tag{6.4}$$

$$I_{n_i} = \frac{K}{\varpi_2} (x_i - \bar{c}_{imax})^{-\frac{2}{m-1}} \tag{6.5}$$

burada $T_{n_{ij}}$ ve I_{n_i}, i noktasının doğru ve belirsizlik üyelik değerleridir ve küme merkezi de c_j'dir. $\bar{c}_{imax}$, $T_{n_{ij}}$'nin en büyük ilk iki değerinin indeksi kullanılarak elde edilmektedir. Onlar her bir yinelemede, ε sonlandırma kriteri olmak üzere, $\left| {T_{n_{ij}}}^{(k+1)} - {T_{n_{ij}}}^{(k)} \right| < \varepsilon$ olana kadar güncellenmektedir.

6.2.2. Belirsizlik Filtrelemesi

Belirsizliğe dayalı yeni bir filtre tanımlanmıştır ve bölütlemedeki belirsizlik bilgisinin etkisini ortadan kaldırmak için kullanılmıştır. Çekirdek fonksiyonu aşağıdaki gibi bir Gauss fonksiyonu kullanılarak tanımlanabilir:

$$G_I(u,\mathrm{v}) = \frac{1}{2\pi\sigma_I^2}\exp\left(-\frac{u^2+v^2}{2\sigma_I^2}\right) \tag{6.6}$$

$$\sigma_I(x,y) = f\big(I(x,y)\big) = aI(x,y) + b \tag{6.7}$$

burada σ_I belirsizlik derecesine bağlı bir fonksiyon $f(\cdot)$ olarak tanımlanan standart sapma değeridir. Belirsizlik seviyesi yüksek olduğunda σ_I büyüktür ve filtreleme mevcut yerel komşulukları daha düzgün hale getirebilir. Belirsizlik seviyesi düşük olduğunda σ_I küçüktür ve filtreleme yerel komşuluklar üzerinde daha az düzgünleştirme işlemi gerektirir. Gauss fonksiyonunu kullanmanın nedeni, belirsizlik derecesini daha düzgün bir filtre ağırlıklarına eşleyebilmesidir (map etme).

$T_s(x,y)$ üzerinde belirsiz bir filtreleme yapılır ve daha homojen hale gelir:

$$\begin{aligned} T_S'(x,y) &= T_S(x,y) \oplus G_{Is}(u,v) \\ &= \sum_{v=y-m/2}^{y+m/2}\sum_{u=x-m/2}^{x+m/2} T_S(x-u,y-v)G_{Is}(\mathrm{u},\mathrm{v}) \end{aligned} \tag{6.8}$$

$$G_{Is}(u,\mathrm{v}) = \frac{1}{2\pi\sigma_{Is}^2}\exp\left(-\frac{u^2+v^2}{2\sigma_{Is}^2}\right) \tag{6.9}$$

$$\sigma_{Is}(x,y) = f\big(I_S(x,y)\big) = aI_S(x,y) + b \tag{6.10}$$

burada T_S' belirsiz filtreleme sonucudur. a ve b doğrusal fonksiyonda belirsizlik seviyesini bir parametre değerine çeviren parametrelerdir. Filtreleme, NCO'dan sonra $T_{n_{ij}}(x,y)$ üzerinde de kullanılır. NCO'nun girişi, belirsizlik filtrelemesinin ardından yerel uzamsal nötrozofik değerdir.

$$T_{n_{ij}}'(x,y) = T_{n_{ij}}(x,y) \oplus G_{In}(u,v)$$

$$= \sum_{v=y-m/2}^{y+m/2} \sum_{u=x-m/2}^{x+m/2} T_{n_{ij}}(x-u, y-v) G_{In}(u,v) \tag{6.11}$$

$$G_{In}(u,\mathrm{v}) = \frac{1}{2\pi\sigma_{In}^2} \exp\left(-\frac{u^2+v^2}{2\sigma_{In}^2}\right) \tag{6.12}$$

$$\sigma_{In}(x,y) = f\big(I_n(x,y)\big) = cI_n(x,y) + d \tag{6.13}$$

burada $T_{n_{ij}}'$, T_S üzerindeki belirsiz filtreleme sonucudur ve m filtre çekirdeğinin boyutudur. Bir çizge oluşturmak için $T_{n_{ij}}'$ ve görüntüyü bölütlemek için en büyük akış algoritması kullanılmıştır.

6.2.3. NÇK: Nötrozofik Çizge Kesim Yöntemi

Bir $C = (S,T)$ kesimi, bir $G = (V,E)$ çizgesini iki alt kümeye ayırır: S ve T. Bir $C = (S,T)$ kesiminin kesim kümeleri bir bitiş noktası S'de ve diğer bitiş noktası T'de olan kenarların $\{(u,v) \in E | u \in S, v \in T\}$ kümesidir. Çizge kesimler, bir çizgeyi en büyük akış problemine veya çizgenin en az kesimine dönüştüren, enerjiyi en aza indirgeme açısından formüle ederek görüntü bölütleme problemlerini etkili bir şekilde çözebilir.

Enerji fonksiyonu genellikle iki bileşenden oluşmaktadır: veri kısıtı E_{veri} ve düzgünlük kısıtı $E_{d\ddot{u}zg\ddot{u}n}$.

$$E(f) = E_{veri}(f) + E_{d\ddot{u}zg\ddot{u}n}(f) \tag{6.14}$$

burada f pikselleri farklı gruplara atayan bir haritadır (eşleme). E_{veri}, f ve atanmış gölge arasındaki benzeşmezliği ölçüp bunu t-link olarak gösterirken, $E_{d\ddot{u}zg\ddot{u}n}$, f 'nin parça düzgünlüğünün kapsamını değerlendirir ve bir çizgede n-link olarak göstermektedir.

Enerji fonksiyonu uygulamasında farklı modellerin farklı biçimleri vardır. Potts modeline dayanan fonksiyon şu şekilde tanımlanmaktadır:

$$E(f) = \sum_{p \in P} D_p\,(f_p) + \sum_{\{p,q\} \in N} V_{\{p,q\}}\,(f_p, f_q) \tag{6.15}$$

burada p ve q pikseller ve N ise p'nin komşuluğudur. D_p, p pikseli için bir bölütlemenin ne kadar uygun olduğunu değerlendirmektedir.

Önerilen NÇK algoritmasında, veri fonksiyonu D_p ve düzgünlük fonksiyonu $V_{\{p,q\}}$ şu şekilde tanımlanmıştır:

$$D_{ij}(p) = \left| T_{n_{ij}}'(p) - C_j \right| \tag{6.16}$$

$$V_{\{p,q\}}(f_p, f_q) = u\delta(f_p \neq f_q) \tag{6.17}$$

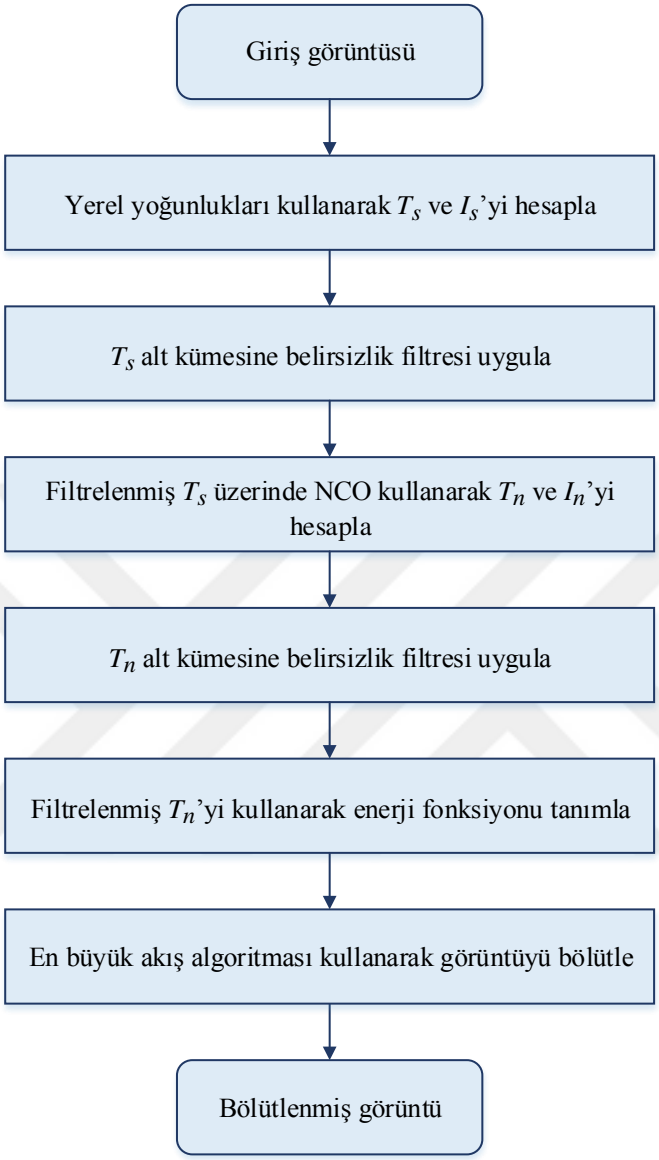
$$\delta(f_p \neq f_q) = \begin{cases} 1\ eğer\ (f_p \neq f_q) \\ 0\ diğer\ durumlarda \end{cases} \tag{6.18}$$

burada u, $[0,1]$'deki sabit bir sayıdır ve p ve q piksellerinin etiketlemesindeki anlaşmazlığın bir cezası olarak kullanılmaktadır.

Enerji fonksiyonu, NK alanında yeniden tanımlandıktan sonra nesneleri arka plandan ayırmak için ÇK teorisinde en büyük akış algoritması kullanılmıştır. Bütün adımlar şu şekilde özetlenebilir:

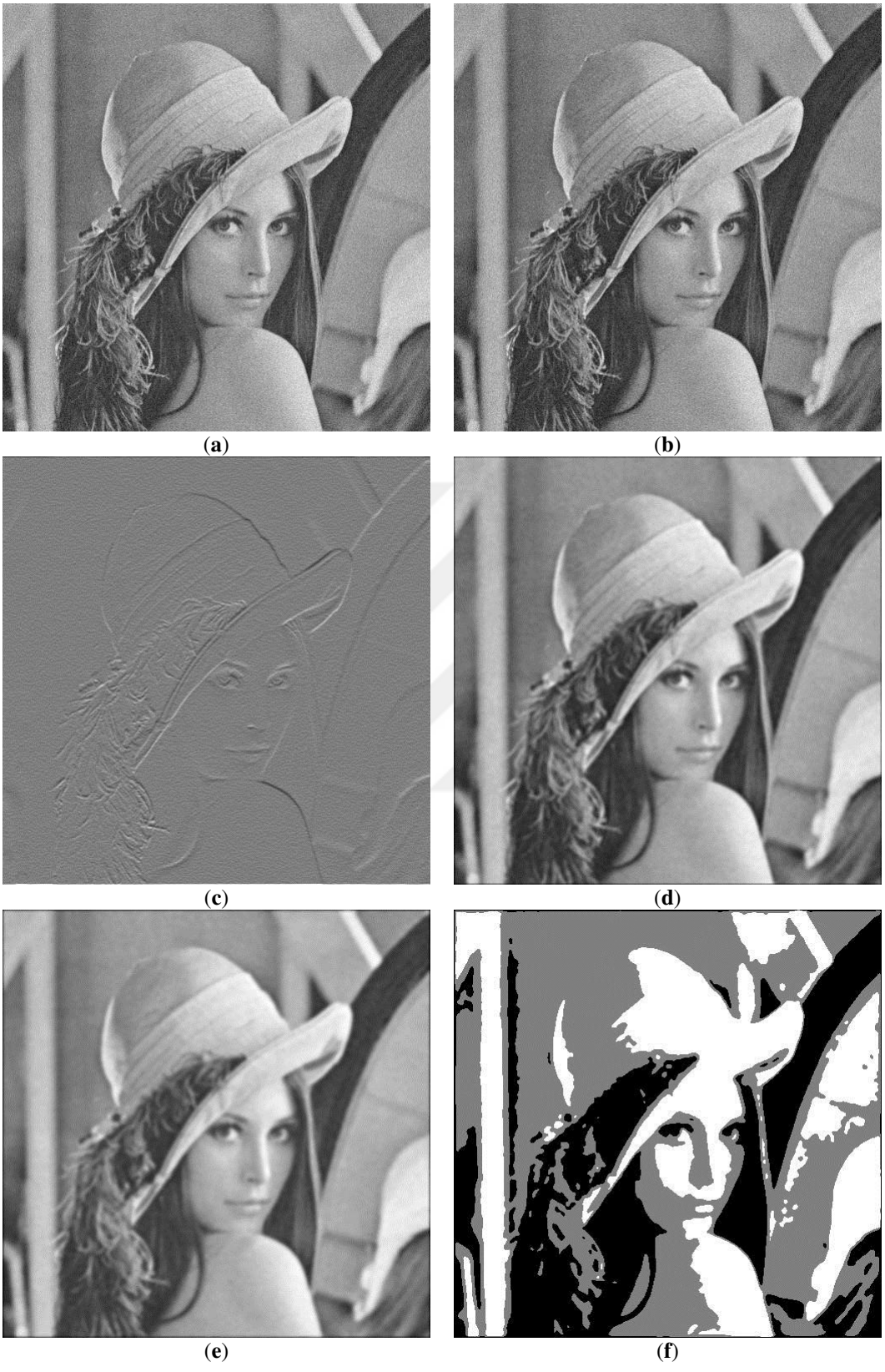
Adım 1:	Yerel nötrozofik değerler T_S ve I_S'yi hesapla;
Adım 2:	I_S'yi kullanarak T_S üzerinde belirsizlik filtrelemesi yap;
Adım 3:	T_n ve I_n'yi elde etmek için filtrelenmiş T_S alt kümesi üzerinde NCO algoritması kullan;
Adım 4:	I_n'ye dayalı belirsizlik filtresi kullanarak T_n'yi filtrele;
Adım 5:	T_n değerine dayalı enerji fonksiyonunu hesapla;
Adım 6:	En büyük akış algoritması kullanarak görüntüyü bölütle.

Önerilen yaklaşımın akış şeması Şekil 6.1'de gösterilmiştir.



Şekil 6.1. Önerilen NÇK yönteminin akış şeması

Tüm algoritmanın adımlarını göstermek için, Şekil 6.2'de örnek bir görüntü kullanarak bazı ara sonuçlar gösterilmiştir.



(**a**) (**b**) (**c**) (**d**) (**e**) (**f**)

Şekil 6.2. "*Lena*" görüntüsü için ara sonuçlar: (a) Orijinal görüntü, (b) T_s sonucu, (c) I_s sonucu, (d) Filtrelenmiş T_s sonucu, (e) T_n'nin filtreleme sonucu, (f) Nihai sonuç

6.3. Deneysel Sonuçlar

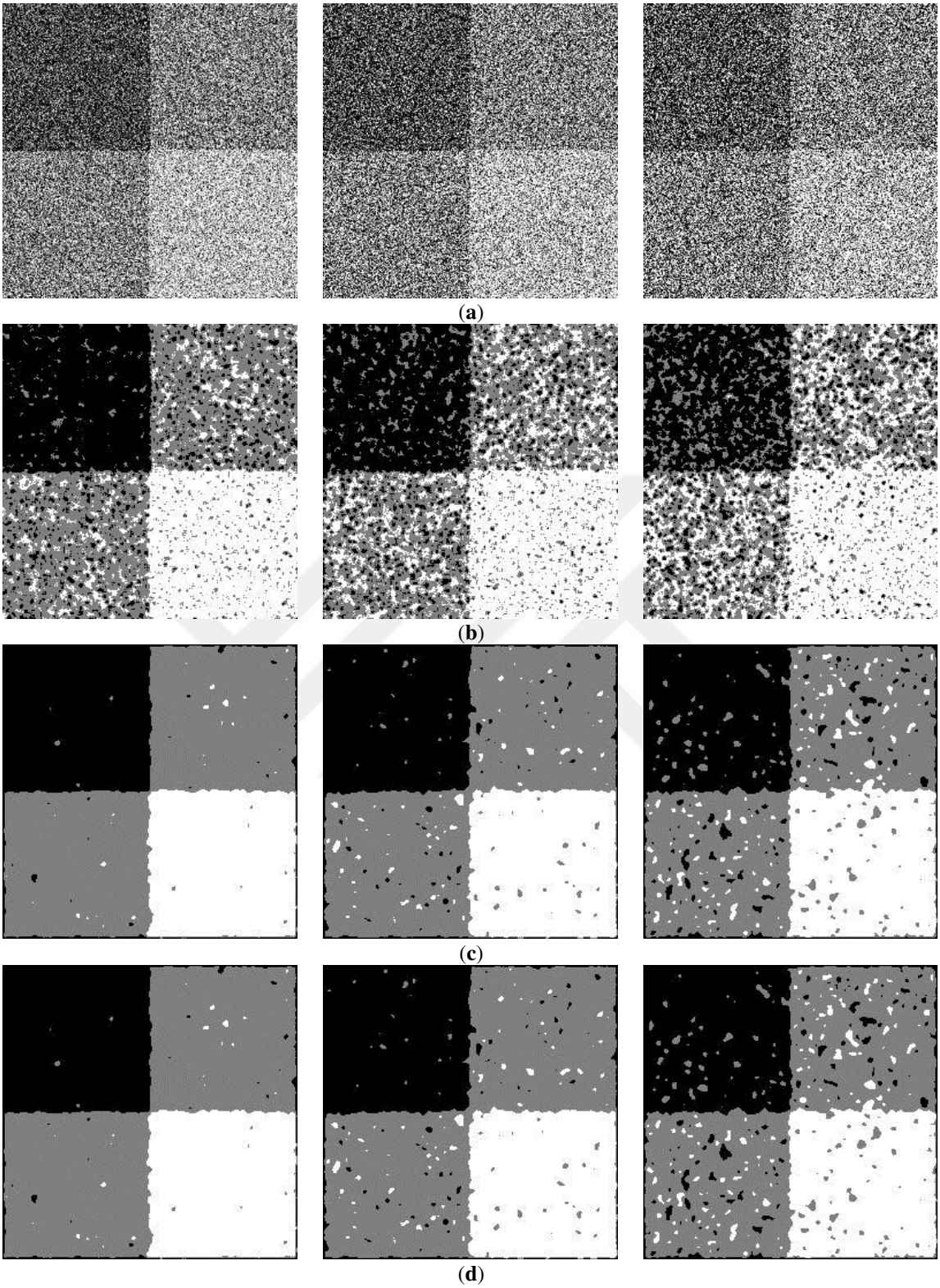
Gürültü gibi belirsiz bilgiye sahip görüntüleri bölütlemek zordur. Bu problemi çözmek için farklı algoritmalar geliştirilmiştir. NÇK yaklaşımı, görüntü bölütlemedeki başarım değerlendirmesi için birçok görüntü üzerinde test edilmiştir. Önceki yöntemlerden [68] daha iyi başarım gösteren son zamanlarda yayınlanmış nötrozofik benzerlik kümeleme (NBK) yöntemi [77] ve yeni geliştirilen bir ÇK yöntemi ile karşılaştırılmıştır [118].

Tüm deneyler aynı parametre değerleri kullanılarak yapılmıştır: $a = 10$; $b = 0{,}25$; $c = 10$; $d = 0{,}25$ ve $u = 0{,}5$.

6.3.1. Niceliksel Değerlendirme

NÇK yöntemini, NBK ve ÇK yöntemleriyle görsel olarak karşılaştırmak için yapay gürültülü görüntüler kullanılmıştır. Sonrasında başarımları iki ölçüt kullanarak niceliksel olarak test edilmiştir. NBK yönteminde [77], başarımı değerlendirmek için yapay gürültülü görüntüler kullanılmıştır. Karşılaştırmanın adil ve tutarlı olması için aynı görüntü ve gürültü seçilmiş ve üç algoritma da bunlar üzerinde test edilmiştir.

64, 128 ve 192 yoğunluklarına sahip yapay bir görüntüye Gauss gürültüsü eklenmiş ve bu görüntü NÇK, NBK ve ÇK algoritmalarının başarımını değerlendirmek için kullanılmıştır. Şekil 6.3 (a)'da gürültü ortalama değeri 0 ve varyans değerleri sırasıyla 80, 100 ve 120 olan orijinal gürültülü görüntü gösterilmiştir. Şekil 6.3 (b)-(d)'de sırasıyla NBK, ÇK ve NÇK yöntemlerine göre sonuçlar listelenmiştir. Ayrıca Şekil 6.3'teki sonuçlar NÇK yönteminin düşük kontrastlı ve gürültülü yapay görüntülerde NBK ve ÇK yöntemlerinden görsel olarak daha iyi başarım gösterdiğini ortaya koşmuştur. Şekil 6.3 (b) ve (c)'de yanlış gruplara atanmış pikseller, Şekil 6.3 (d)'de NÇK yöntemiyle doğru gruplara atanmıştır. Etiketlemeyi zorlaştıran sınır pikselleri de NÇK tarafından doğru kategorilere bölütlenmiştir.



Şekil 6.3. Düşük kontrastlı ve gürültülü yapay bir görüntü üzerinde bölütleme karşılaştırması (a) Farklı seviyelerde Gauss gürültüsü olan yapay görüntü, (b) NBK sonuçları, (c) ÇK sonuçları, (d) NÇK sonuçları

Yanlış sınıflandırma hatası (YSH), bölütleme başarımlarını değerlendirmek için kullanılmaktadır [119–121]. YSH, yanlışlıkla ön planda kategorize edilmiş arka planın yüzdesini ve bu durumun tersini ölçmektedir.

$$YSH = 1 - \frac{|B_O \cap B_T| + |F_O \cap F_T|}{|B_O| + |F_O|} \tag{6.19}$$

burada F_O, B_O, F_T ve B_T sırasıyla kesin-referans görüntüsü ve elde edilen sonuç görüntüsü üzerindeki nesne (ön plan) ve arka plan pikselleridir.

Buna ek olarak, Başarım ölçüsü (BÖ) (Figure of Merit)(FOM) [120], bölütlenmiş sonuç ile kesin-referans (ground truth) arasındaki farkı değerlendirmek için kullanılmıştır:

$$B\ddot{O} = \frac{1}{\max(N_I, N_A)} \sum_{k=1}^{N_A} \frac{1}{1 + \beta d^2(k)} \tag{6.20}$$

burada N_I ve N_A, bölütlenmiş nesnenin ve doğru nesnelerin piksellerinin sayısıdır. $d(k)$, k'nıncı gerçek pikselden en yakın bölütlenmiş sonuç pikseline olan uzaklıktır. β, bir sabittir ve kaynak [31]'de $1/9$ olarak seçilmiştir.

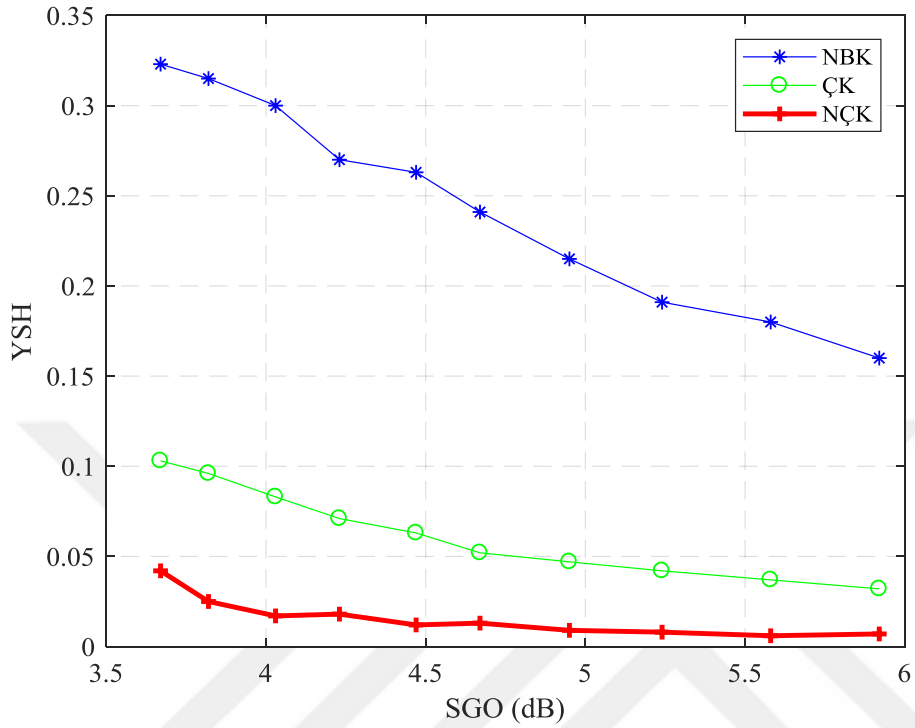
Gürültülü görüntünün kalitesi bir sinyal gürültü oranı (SGO) ile ölçülmektedir:

$$SGO = 10\log\left[\frac{\sum_{r=1}^{H-1}\sum_{c=1}^{W-1} I^2(r,c)}{\sum_{r=1}^{H-1}\sum_{c=1}^{W-1}\left(I(r,c) - I_n(r,c)\right)^2}\right] \tag{6.21}$$

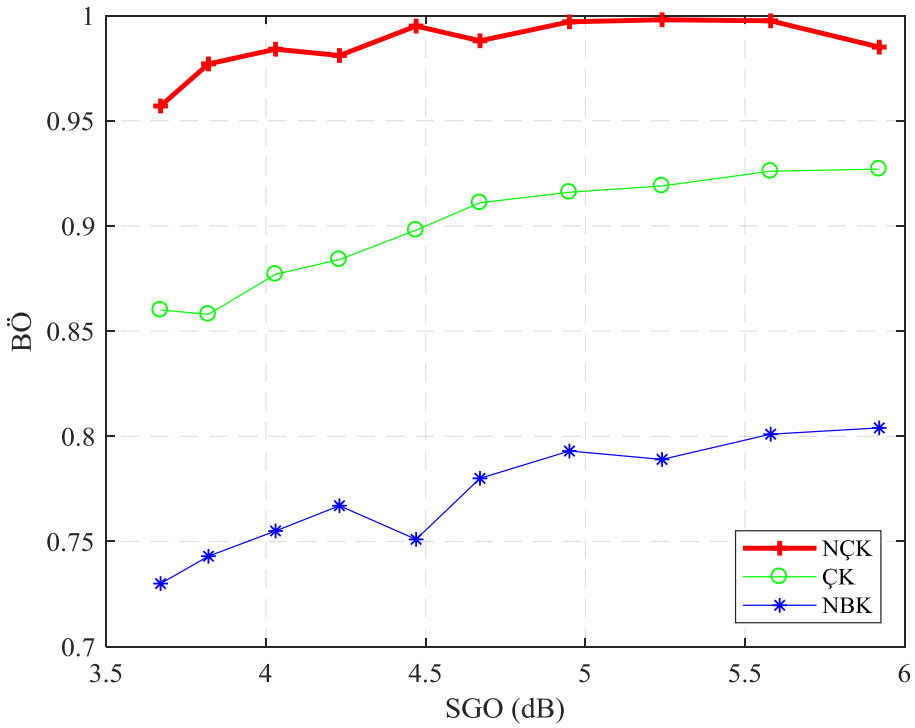
burada $I_n(r,c)$ ve $I(r,c)$ sırasıyla gürültülü ve orijinal görüntüdeki (r,c) noktasının yoğunluk değerleridir.

YSH ve BÖ sonuçları Şekil 6.4 ve 6.5'te çizilmiştir; burada "$*$" NBK yöntemini, "o" ÇK yöntemini ve "$+$" NÇK yöntemini göstermektedir. NÇK yöntemi en düşük YSH değerlerine sahiptir. NÇK'nin tüm YSH değerleri 0,043'ten daha küçüktür, NBK ve ÇK yöntemlerinden elde edilen tüm değerler NÇK yönteminin değerlerinden daha büyüktür. NBK'nin en düşük YSH değeri 0,1614 ve ÇK'nin en düşük YSH değeri 0,0327 iken; NÇK

SGO 5,89 dB olduğunda YSH=0,0068 değeriyle en iyi başarımı elde etmiştir. NÇK özellikle düşük SGO'da, NBK ve ÇK'dan daha büyük BÖ değerlerine sahiptir.



Şekil 6.4. YSH'nin çizimi ∗: NBK yöntemi, o: ÇK yöntemi, +: NÇK yöntemi



Şekil 6.5. BÖ'nün çizimi ∗: NBK yöntemi, o: ÇK yöntemi, +: NÇK yöntemi

Karşılaştırma sonuçları Tablo 6.1'de listelenmiştir. YSH ve BÖ'nün sırasıyla ortalama ve standart sapması NBK yöntemi kullanıldığında 0,247 ± 0,058 ve 0,771 ± 0,0252; ÇK yöntemi kullanıldığında 0,062 ± 0,025 ve 0,897 ± 0,027; NÇK kullanıldığında 0,015 ± 0,011 ve 0,987 ± 0,012'dir. NÇK yöntemi daha düşük YSH ve BÖ değerleriyle NBK ve ÇK yönteminden daha iyi başarım elde etmiştir.

Tablo 6.1. Değerlendirme ölçütleriyle başarım karşılaştırmaları

Ölçütler	NBK	ÇK	NÇK
YSH	0,247 ± 0,058	0,062 ± 0,025	0,015 ± 0,011
BÖ	0,771 ± 0,025	0,897 ± 0,027	0,987 ± 0,012

6.3.2. Doğal Görüntülerde Başarım

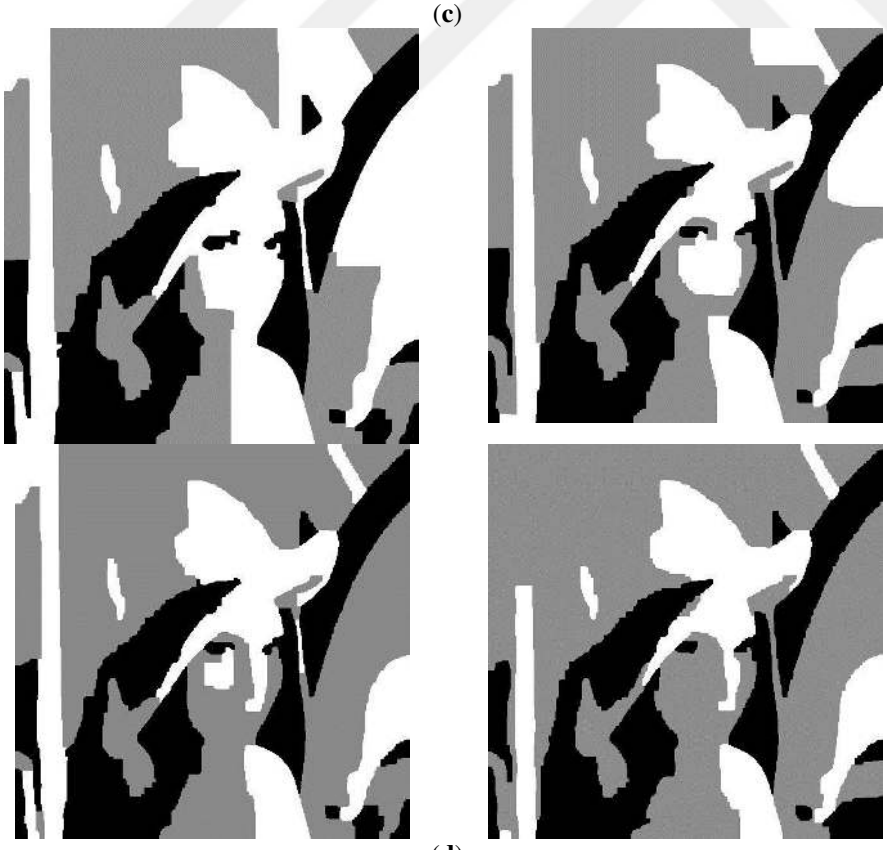
NÇK'nın başarımını doğrulamak için birçok görüntü kullanılmıştır. Ayrıca iyileştirilmiş bir çekirdek çizge kesim (ÇÇK) algoritmasına [122] dayanan yeni geliştirilmiş bir görüntü bölütleme algoritması ile sonuçlar karşılaştırılmıştır. Burada, NÇK yönteminin bölütleme başarımını göstermek için beş tane görüntü rastgele seçilmiştir. Şekil 6.6-6.10'daki ilk satır orijinal görüntüyü ve sırasıyla NBK, ÇK, ÇÇK ve NÇK'nın bölütleme sonuçlarını göstermektedir. Diğer satırlar gürültülü görüntülerin sonuçlarını göstermektedir. NÇK'nın sonuçları NBK, ÇK ve ÇÇK'dan görsel olarak daha iyi kalitededir. Orijinal görüntülerde, NÇK ve ÇK benzer şekilde doğru sonuçlar elde ederken, ÇGC yetersiz bölütleme elde etmiştir. Gürültü arttığı zaman NBK ve ÇK bundan oldukça etkilenir ve aşırı bölütlemeye maruz kalırlar. ÇÇK sonuçları da yetersiz bölütleme ile bazı ayrıntıları kaybeder. Bununla birlikte, NÇK gürültüye maruz kalmaz ve çoğu piksel doğru gruplara ayrılır; sınırdaki ayrıntılar iyice bölütlenmiştir.

Şekil 6.6 "*Lena*" görüntüsündeki bölütleme sonuçlarını göstermektedir. Şeklin son satırında bulunan NÇK diğerlerinden daha iyi sonuçlar vermiştir. Yüz, burun, ağız ve göz bölgeleri NÇK tarafından doğru bir şekilde bölütlenmiştir. Saç bölgesi ve şapkanın üstündeki alan gibi gürültülü bölgeler de doğru bir şekilde bölütlenmiştir. Bunula birlikte, NBK ve ÇK yöntemleri özellikle şapkanın üstündeki bölgede yanlış bölütleme elde etmiştir. ÇÇK sonuçları yüzdeki ve gözlerdeki bazı detay bilgileri kaybetmiştir.

(**a**)

(**b**)

(**c**)



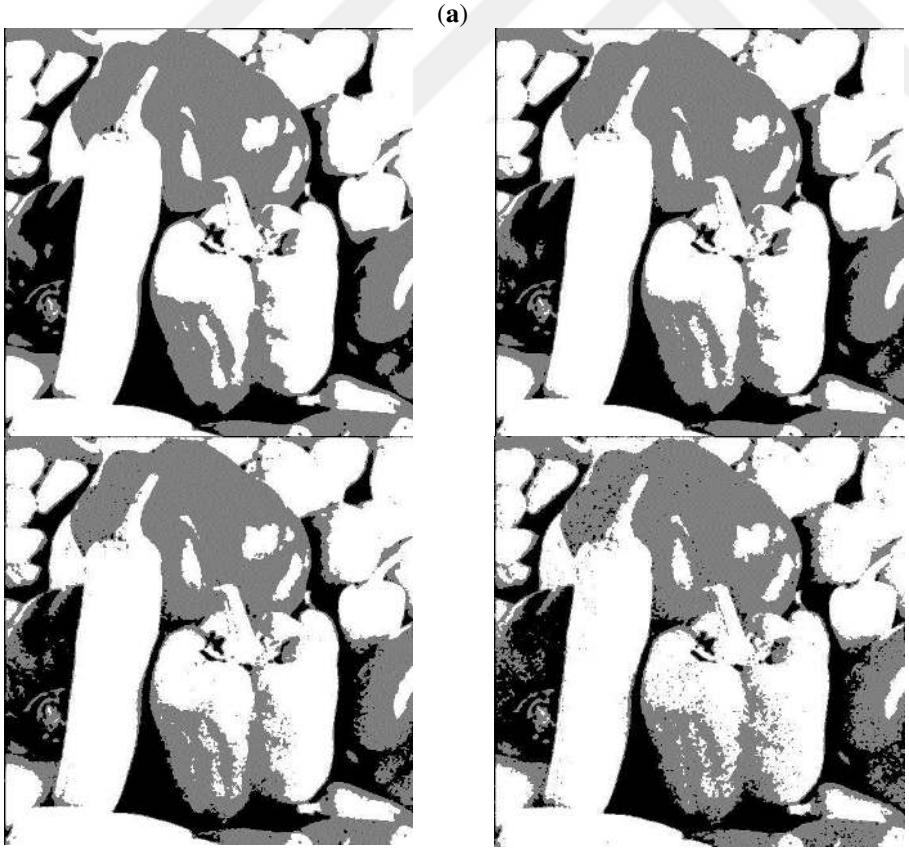
(**d**)

(**e**)

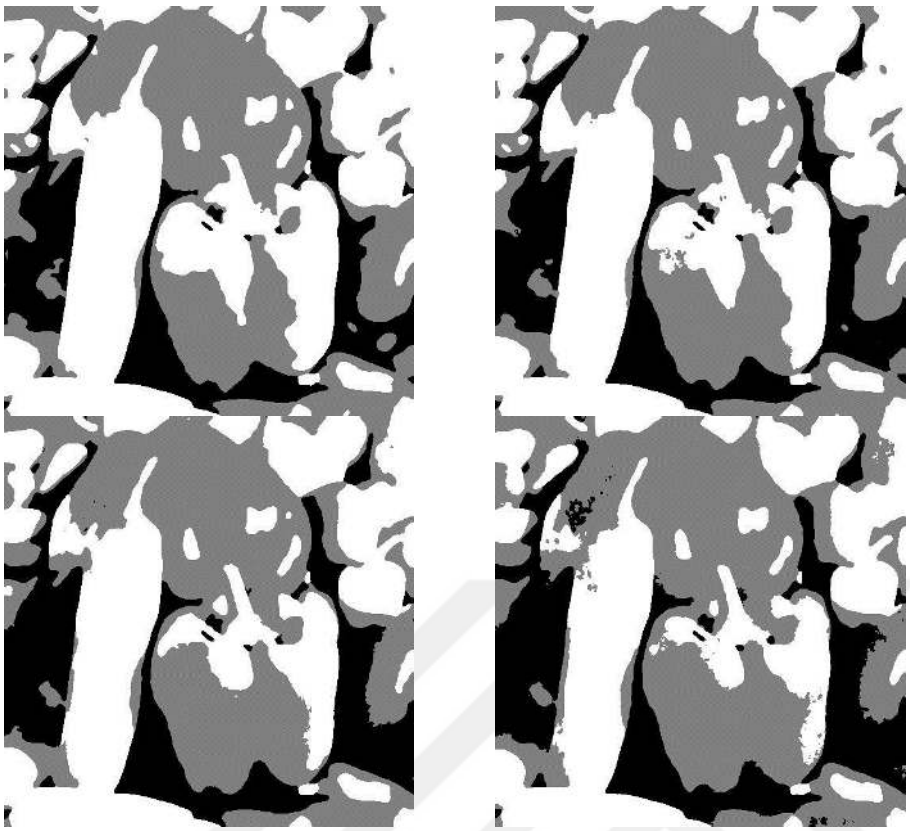
Şekil 6.6. "*Lena*" görüntüsündeki karşılaştırma sonuçları (a) Farklı Gauss gürültü seviyesine sahip "*Lena*" görüntüsü, varyans: 0, 10, 20, 30, (b) NBK bölütleme sonuçları, (c) ÇK bölütleme sonuçları, (d) ÇÇK bölütleme sonuçları, (e) NÇK bölütleme sonuçları

Şekil 6.7'de gösterildiği gibi "*Peppers*" görüntüsü üzerinde tüm yöntemlerin başarımları karşılaştırılmıştır. Diğer karşılaştırmalar için daha önce de belirtildiği gibi sıfır gürültü seviyesi için ÇK, NÇK ve ÇÇK benzer bölütlemeler üretmiştir. ÇK, ÇÇK ve NÇK yöntemleri tüm gürültü seviyelerinde NBK'dan daha iyi bölütleme sonuçları elde etmiştir. Gürültü seviyesi artıkça önerilen NÇK yönteminin etkinliği daha belirgin hale gelmiştir. ÇK sonuçlarında bazı yanlış bölütleme bölgeleri (gri biber bölgelerinde siyah bölgeler) vardır. Arka plan bölgelerinin bazıları da ÇK yöntemi ile yanlış bölütlere ayrılmıştır. Önerilen NÇK yöntemi ile daha uygun ve daha doğru bölütler elde edilmiştir. Özellikle gürültü seviyeleri 20 ve 30 için daha az yanlış bölütleme bölgeleri ürettiğinden, NÇK yönteminin bölütleme başarısı görsel olarak diğerlerinden daha iyidir. Bu görüntü üzerinde ÇÇK, bölütleme sonuçları bakımından NÇK ile benzer bir başarım elde etmiştir.

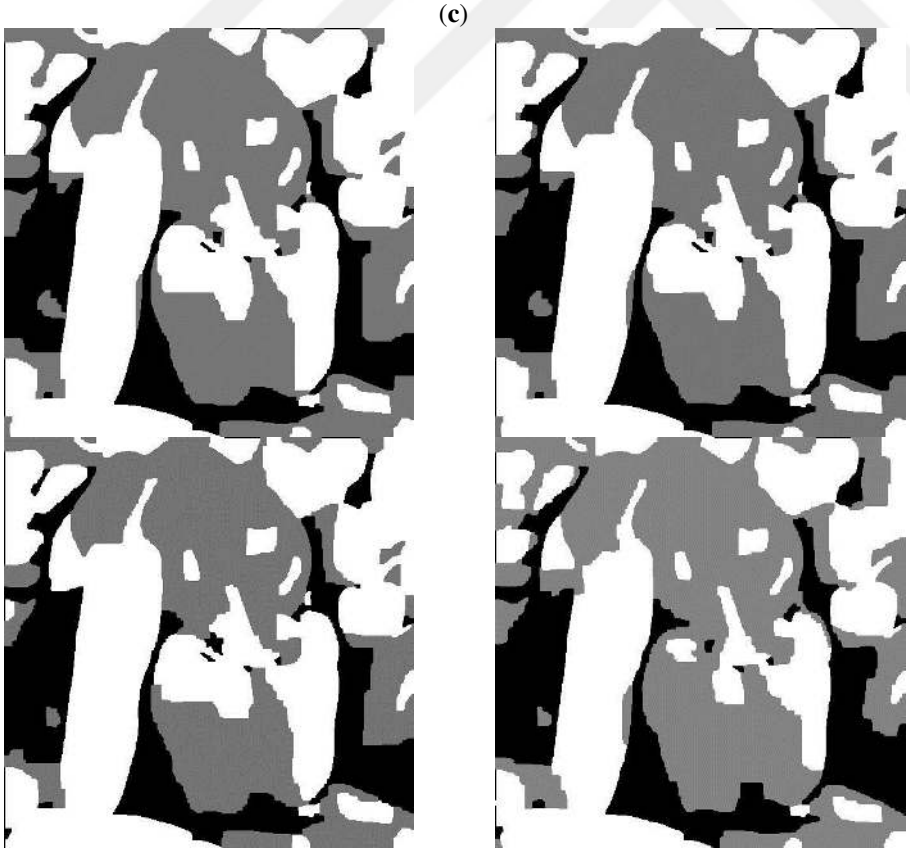
(**a**)



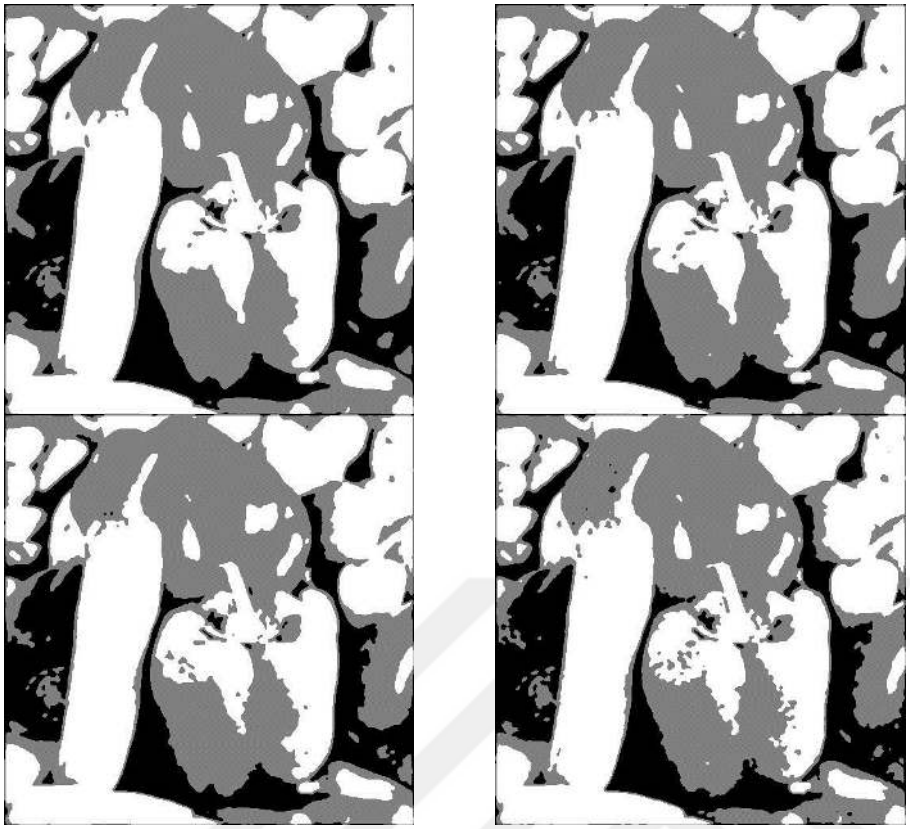
(**b**)



(**c**)



(**d**)



(**e**)

Şekil 6.7. "*Peppers*" görüntüsündeki karşılaştırma sonuçları (a) Farklı Gauss gürültü seviyesine sahip "*Peppers*" görüntüsü, varyans: 0, 10, 20, 30, (b) NBK bölütleme sonuçları, (c) ÇK bölütleme sonuçları, (d) ÇÇK bölütleme sonuçları, (e) NÇK bölütleme sonuçları

"*Woman*" görüntüsündeki karşılaştırma sonuçları Şekil 6.8'de verilmiştir. Gürültü seviyesi arttığında NBK yönteminin daha kötü bölütleme sonuçları ürettiği açıktır. Daha homojen bölgeler üreten ÇK ve ÇÇK yöntemleri, NBK yöntemi ile karşılaştırıldığında daha iyi sonuçlar vermiştir. Ayrıca ÇK, ÇÇK ve NBK yöntemlerinin gürültüsüz durumda aynı bölütleme sonuçları ürettiğini söylememiz gerekir. Ancak gürültü seviyesi arttıkça kadının yüzü daha karışık hale gelmektedir. Öte yandan önerilen NÇK yöntemi, diğer yöntemlere kıyasla daha belirgin bölgeler üretmiştir. ÇÇK sonuçlarında gözlerin ve burnun sınırları tanınamaz hale gelmiştir. Bunun yanında NÇK tarafından üretilen bölgelerin kenarları daha düzgün ve pürüzsüzdür.

(**a**)

(**b**)

(**c**)

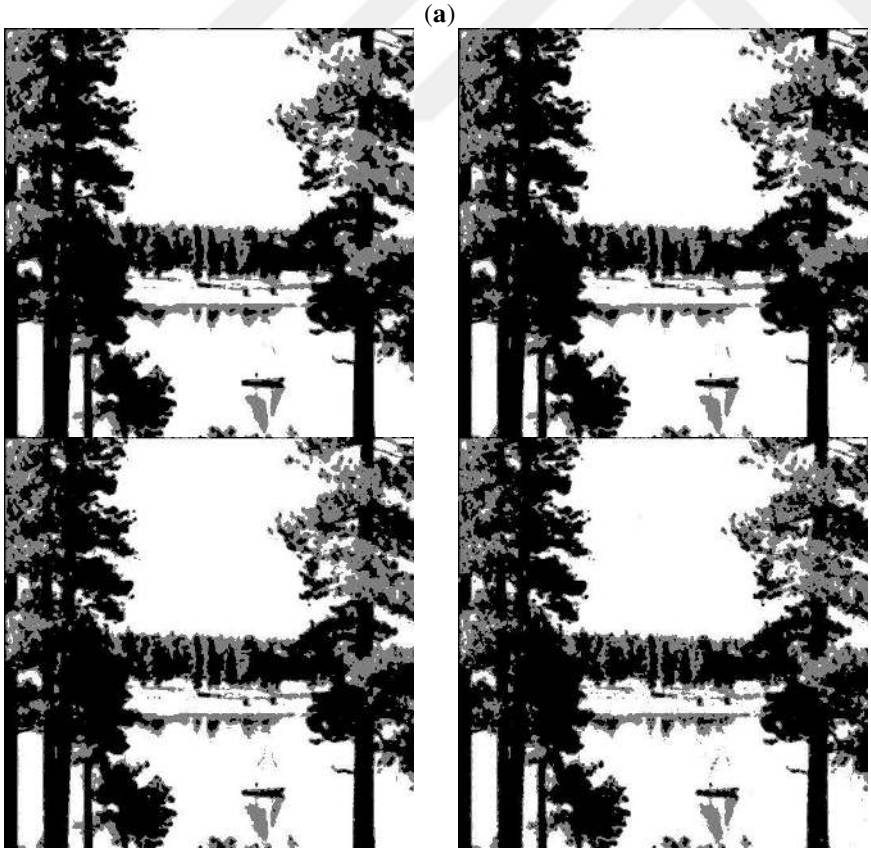
(**d**)

(**e**)

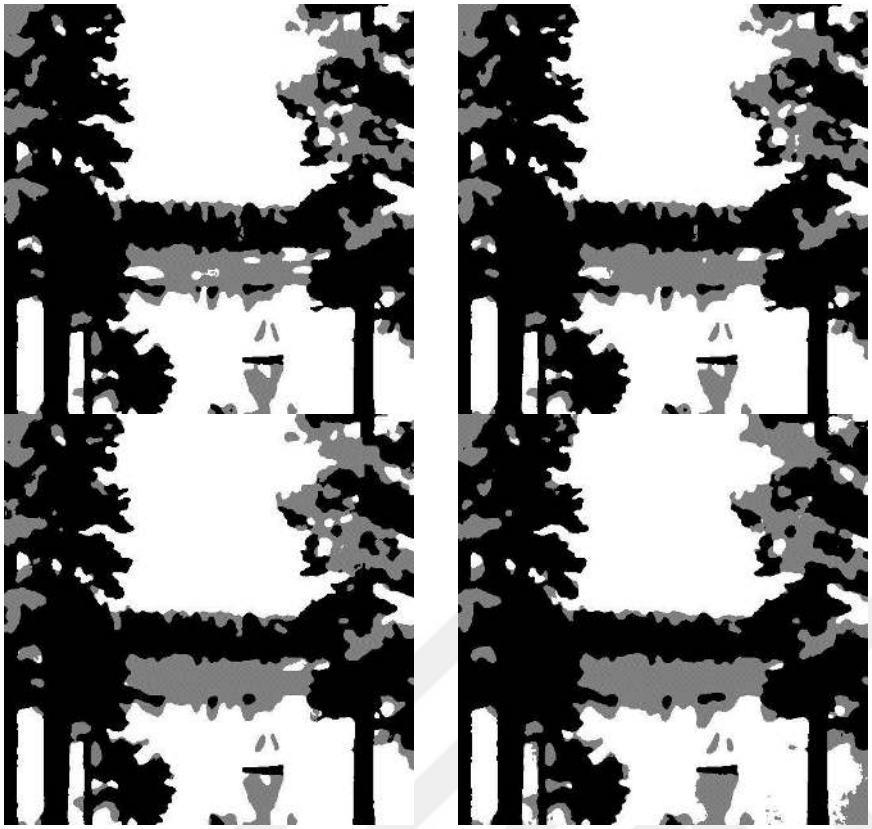
Şekil 6.8. "*Woman*" görüntüsündeki karşılaştırma sonuçları (a) Farklı Gauss gürültü seviyesine sahip "*Woman*" görüntüsü, varyans: 0, 10, 20, 30, (b) NBK bölütleme sonuçları, (c) ÇK bölütleme sonuçları, (d) ÇÇK bölütleme sonuçları, (e) NÇK bölütleme sonuçları

Şekil 6.9'da gösterildiği gibi bu yöntemler "*Lake*" görüntüsüyle de karşılaştırılmıştır. Karşılaştırmalarda, ÇK, ÇÇK ve NÇK yöntemlerinin NBK yönteminden daha iyi sonuçlar ürettiği görülmüştür. Sonuçlar özellikle yüksek gürültü seviyelerinde daha iyidir. ÇK ve ÇÇK yöntemlerinin daha homojen bölgeler ürettiğini belirtmek gerekir, ancak bu durumda sınır bilgisi kaybolmuştur. Bu ÇK yönteminin önemli bir dezavantajıdır. Öte yandan, önerilen NÇK yöntemi kenar bilgilerini korumakla birlikte, kıyaslanabilir benzer homojen bölgeler üretmiştir. Önerilen NÇK yöntemi özellikle yüksek gürültü seviyelerinde daha iyi sonuçlar vermiştir.

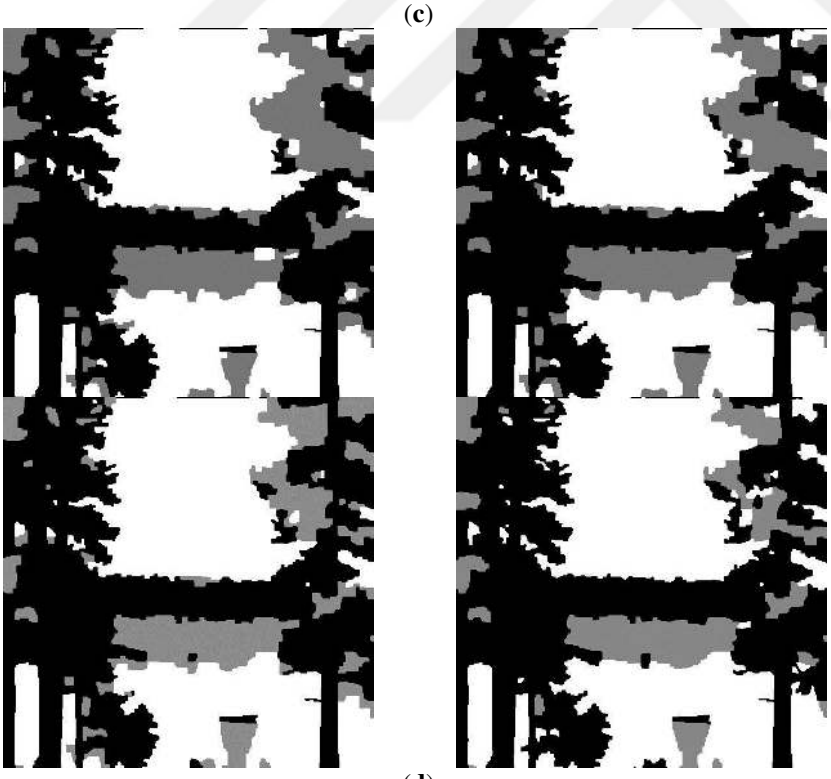
(**a**)



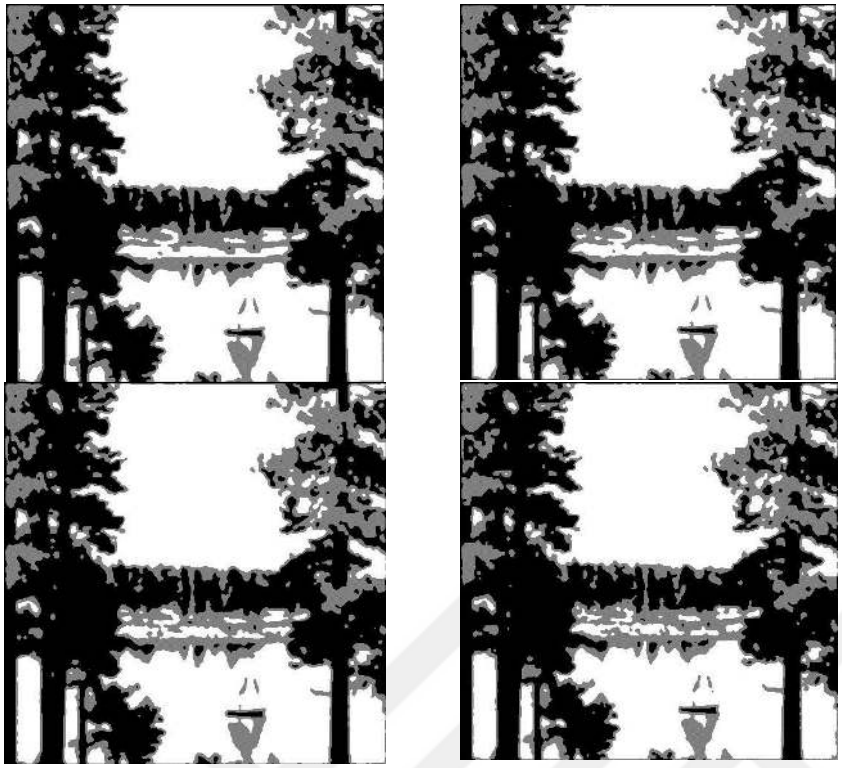
(**b**)



(**c**)



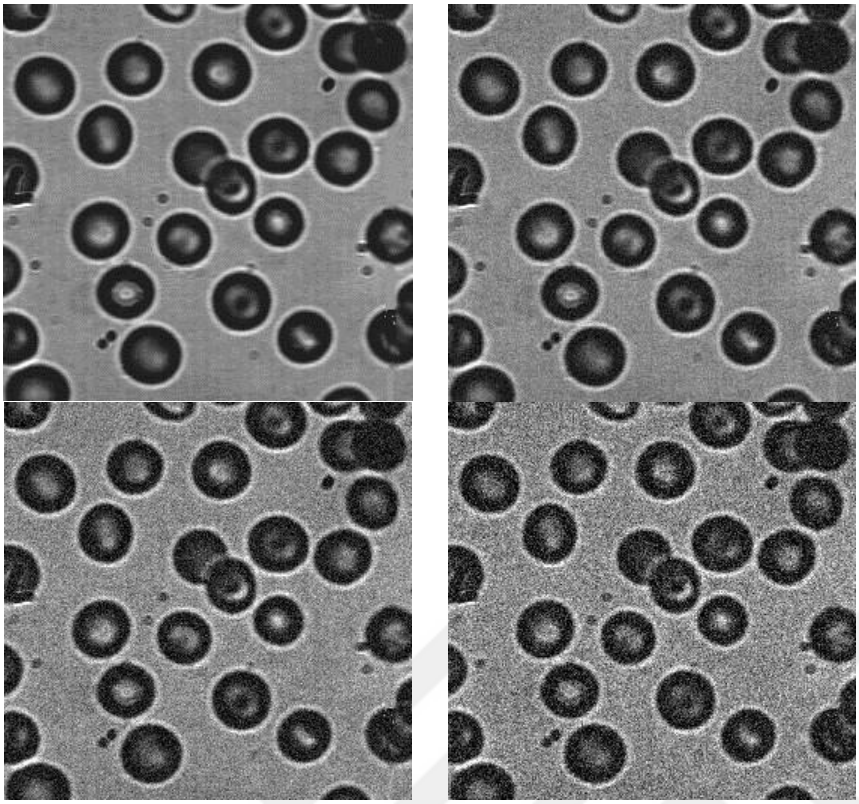
(**d**)



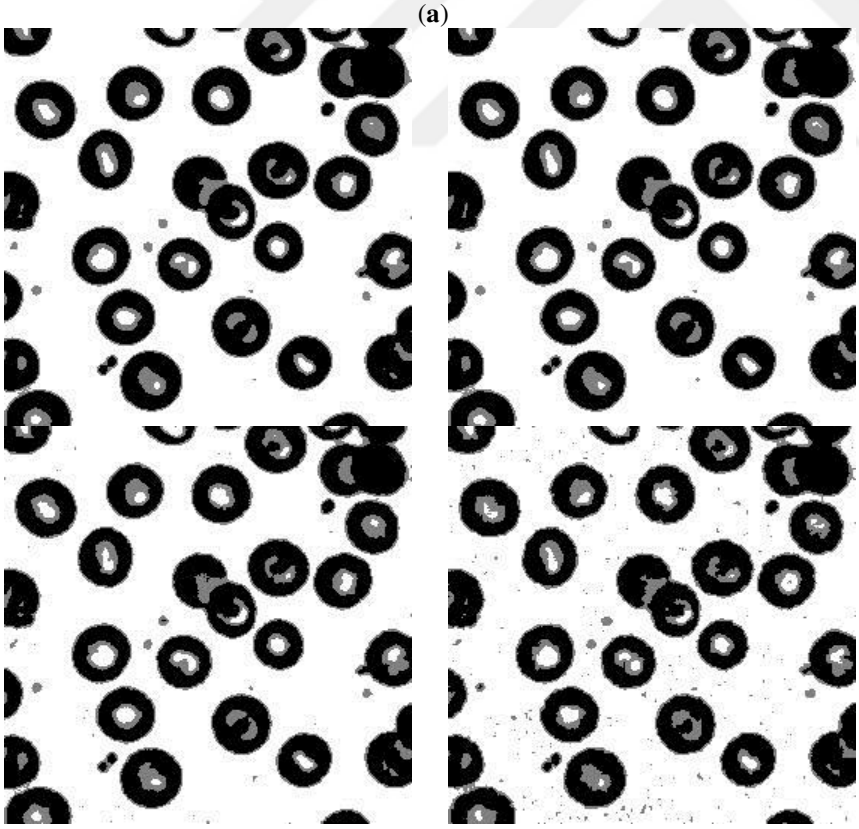
(**e**)

Şekil 6.9. "*Lake*" görüntüsündeki karşılaştırma sonuçları (a) Farklı Gauss gürültü seviyesine sahip "*Lake*" görüntüsü, varyans: 0, 10, 20, 30, (b) NBK bölütleme sonuçları, (c) ÇK bölütleme sonuçları, (d) ÇÇK bölütleme sonuçları, (e) NÇK bölütleme sonuçları

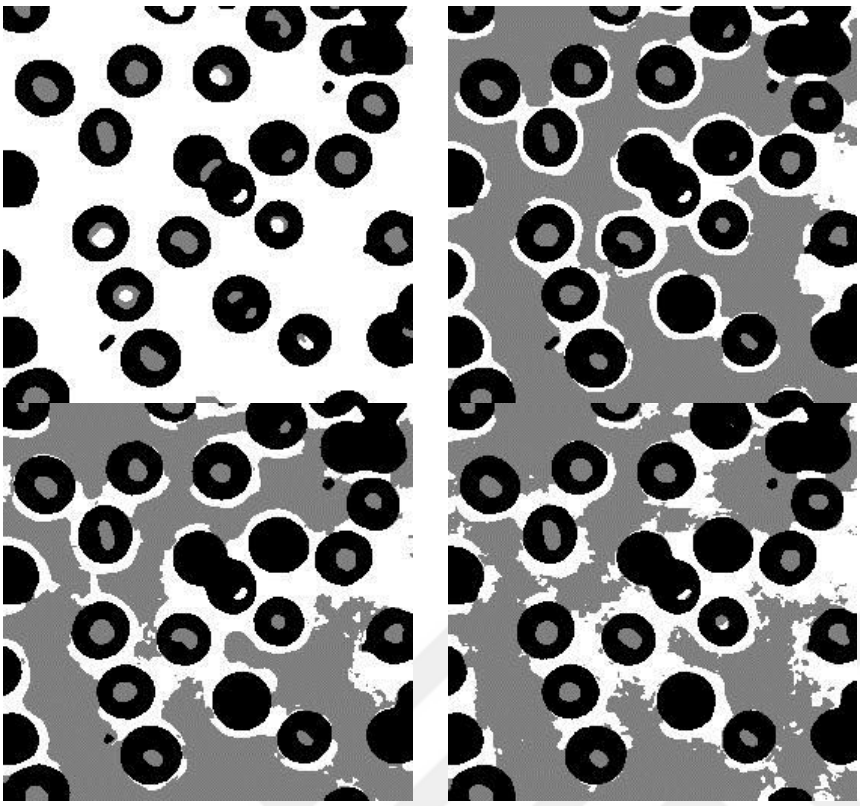
Şekil 6.10'da, karşılaştırma amacıyla daha uygun bir görüntü kullanılmıştır. Kan hücreleri nesne olarak değerlendirilebilirken, görüntünün geri kalanı arka plan olarak düşünülebilir. "*Blood*" görüntüsünde, NBK ve NÇK yöntemleri benzer bölütleme sonuçları üretmiştir. ÇÇK, arka plan bölgesinde bazı yanlış bölütlere ayrılmıştır. NÇK, kan hücrelerinin doğru ve tam olarak çıkarıldığı gürültülü "*Blood*" görüntüsünde daha iyi sonuçlara sahiptir. Bu görüntüde NÇK algoritmasının üstünlüğü de gözlemlenebilir.



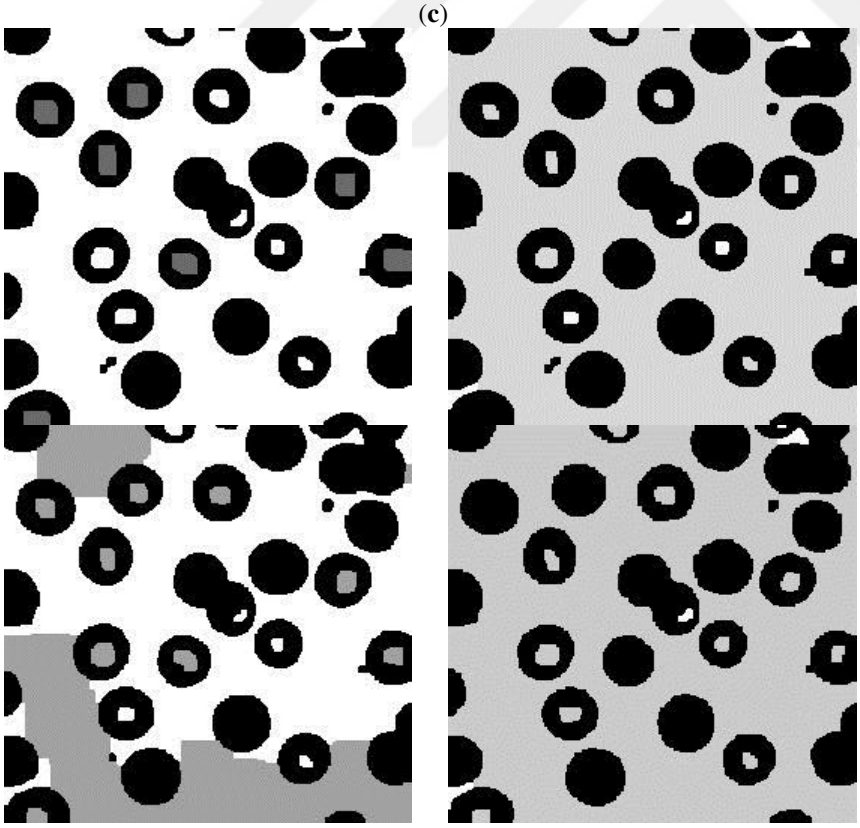
(**a**)



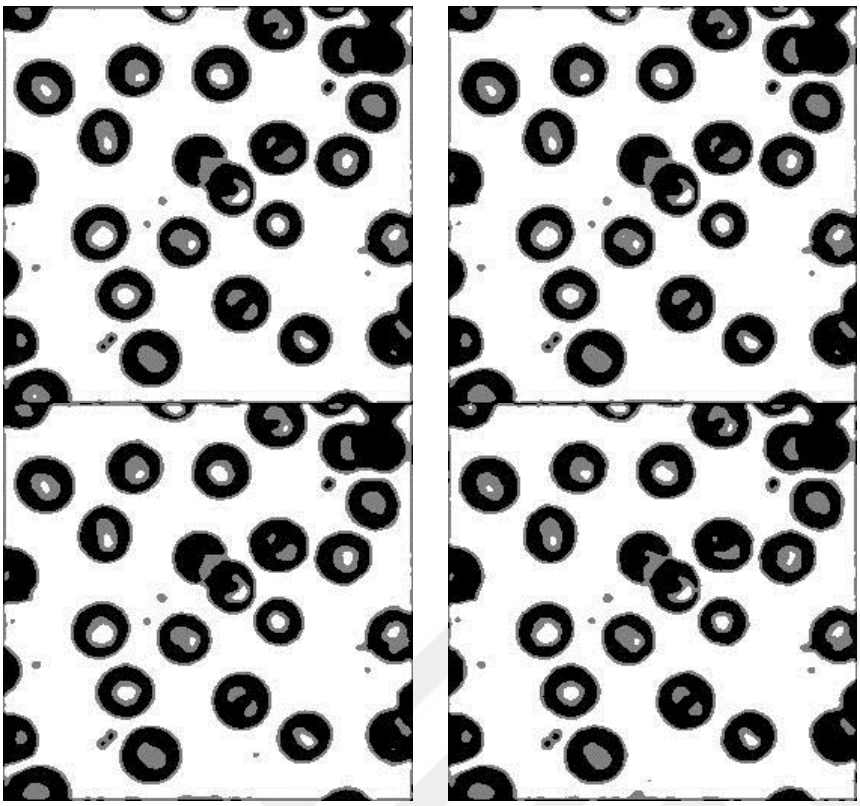
(**b**)



(**c**)



(**d**)



(**e**)

Şekil 6.10. "*Blood*" görüntüsündeki karşılaştırma sonuçları (a) Farklı Gauss gürültü seviyesine sahip "*Blood*" görüntüsü, varyans: 0, 10, 20, 30, (b) NBK bölütleme sonuçları, (c) ÇK bölütleme sonuçları, (d) ÇÇK bölütleme sonuçları, (e) NÇK bölütleme sonuçları

7. SONUÇLAR VE DEĞERLENDİRME

Bu tez çalışmasında, nötrozofik mantık temel alınarak görüntü bölütleme ve örüntü tanıma alanları için yeni algoritmalar ve yeni yöntemler önerilmiştir. Bunlardan ilki, ÇNCO kümeleme algoritmasıdır. Doğrusal olmayan verilerin kümelemesinde etkin olarak kullanılabilen bu algoritma aynı zamanda gürültü ve aykırı veri içeren kümelerle de önemli başarılar elde edebilmektedir. ÇNCO algoritması görüntü bölütleme uygulamalarında iyi sonuçlar vermiştir. İkinci olarak NAAÖM yöntemi önerilmiştir. Bu yöntemde, AÖM nötrozofik kümeye uygun olarak ağırlıklandırılmıştır bunun sonucunda NAAÖM yönteminin dengesiz veri kümeleri üzerinde yüksek sınıflandırma doğruluğuna sahip olduğu görülmüştür. Üçüncü olarak *k*-EYK sınıflandırıcısı nötrozofik açıdan ele alınarak NK-*k*-EYK sınıflandırıcısı önerilmiştir. NK-*k*-EYK sınıflandırıcısının sınıflandırma başarımı artırılması için nötrozofik kümeye dayalı üyelikler benimsenmiştir. Son olarak etkili bir görüntü bölütlemesi için NÇK yöntemi önerilmiştir. Önerilen NÇK yöntemi, nötrozofik belirsizlik filtresine sahip olduğundan hem temiz hem de gürültü gibi belirsizlik bilgisine sahip görüntüleri etkin ve düzgün bir şekilde bölütleyebilmektedir.

7.1. Sonuçların Değerlendirilmesi

Üçüncü bölümde, yeni bir veri kümeleme algoritması olan ÇNCO önerilmiş ve geniş ölçüde deneyler yapılarak verimliliği test edilmiştir. Çekirdek bilgilerinin NCO'ya eklenmesi onu doğrusal olmayan veri kümelemesinde kullanılabilir hale getirmiştir. Önerilen algoritma, çeşitli yapay ve gerçek veri kümelerinin kümelemesinde oldukça başarılıdır. Buna ek olarak, önerilen ÇNCO yönteminin görüntü bölütleme uygulamaları da umut vericidir. Verimliliğin yanı sıra, önerilen ÇNCO yöntemi yeni amaç ve üyelik fonksiyonları nedeniyle gürültü ve aykırı veri noktalarının üstesinden gelmiştir. ÇNCO, belirsizlik bilgisini verimli bir şekilde işleme yeteneğiyle veri madenciliği ve makine öğrenmesi alanlarında daha fazla uygulama alanı bulacaktır.

Dördüncü bölümde, Nötrozofik kümelemeye dayalı yeni bir AAÖM modeli önerilmiştir. Bu yeni ağırlıklandırma şeması, her bir veri noktasının doğru, belirsiz ve yanlış üyeliklerini AÖM'ye tanıtmıştır. Böylece sınıflandırma aşamasında, gürültünün ve aykırı değerlerin etkisini kaldırılabilir ve daha iyi sınıflandırma sonuçlarına ulaşılabilir. Ayrıca, önerilen NAAÖM şeması sınıfsal dengesizlik problemini daha etkin bir şekilde

halletmektedir. Değerlendirme deneylerinde, NAAÖM yönteminin başarımı AÖM, AAÖM, topluluk temelli iki AAÖM ve DVM yöntemleriyle karşılaştırılmıştır. Deneysel sonuçlar, NAAÖM'nin yapay veya gerçek ikili dengesiz veri kümeleri için karşılaştırılan yöntemlerden daha etkili olabileceğini göstermektedir.

Beşinci bölümde, NK-*k*-EYK denilen NK teorisine dayalı yeni bir eğiticili sınıflandırma yöntemi önerilmiştir. Önerilen NK-*k*-EYK yöntemi, üyelikleri eğiticili NCO kümeleme algoritmasına dayalı olarak eğitim örneklerine atamaktadır ve örnekleri nötrozofik üyeliklerine göre sınıflandırmaktadır. Bu yaklaşım, daha önceden önerilmiş bulanık *k*-EYK yönteminin yanlışlık ve belirsizlik kümelerinin eklenmesiyle oluşan bir uzantısı olarak görülebilir. Önerilen yöntemin etkinliği kapsamlı deney sonuçları ile gösterilmiştir. Sonuçlar aynı zamanda diğer geliştirilmiş *k*-EYK yöntemleriyle karşılaştırılmıştır. Elde edilen sonuçlara göre, önerilen NK-*k*-EYK yöntemi çeşitli sınıflandırma uygulamalarında kullanılabilir.

Altıncı bölümde, gürültü gibi belirsiz bilgiye sahip görüntüleri bölütleyebilmek için etkin bir yöntem geliştirilmesi amaçlanmıştır. Bu zorluğun üstesinden gelmek için nötrozofik çizge kesimine dayalı yeni bir görüntü bölütleme yöntemi önerilmiştir. Görüntü, NK alanına eşlenmiş ve yeni tanımlanan bir belirsizlik filtresi kullanılarak filtrelenmiştir. Sonrasında nötrozofik değerlere göre yeni bir enerji fonksiyonu tasarlanmıştır. Belirsizlik filtreleme işlemi, küresel yoğunlukların ve yerel uzamsal bilgilerin belirsizliği ortadan kaldırmıştır. Bölütleme sonuçları en büyük akış algoritması ile elde edilmiştir. Karşılaştırma sonuçları, önerilen yöntemin mevcut yöntemlere göre daha iyi başarım gösterdiğini hem nicel hem de nitel olarak göstermiştir. Ayrıca önerilen NÇK yöntemi, bir görüntüdeki belirsizlik bilgilerini iyi bir şekilde işleyebildiğinden; hem temiz görüntüleri hem de gürültülü görüntüleri düzgün ve etkin bir şekilde bölütlere ayırabildiği görülmüştür.

7.2. Gelecekteki Çalışmalar

Doktora tez sürecinde yapılan çalışmaların ve elde edilen bulguların değerlendirilmesi sonucunda ileri zamanlarda yapılması düşünülen çalışmalar aşağıda belirtilmiştir:

- Gelecekte çalışmalarda NAAÖM yönteminin çok sınıflı dengesiz veri kümelerine uygulanabilmesi için geliştirilmesi düşünülmektedir.
- Bir diğer gelecek çalışma konusu olarak NK-*k*-EYK'nin dengesiz veri kümesi problemlerine uygulanması planlanmaktadır.

- Deney sonuçlarının, Freidman testi ve Wilcoxon signed-rank testi gibi parametrik olmayan istatistiksel yöntemlerle analiz edilmesi amaçlanmaktadır.

7.3. Yayınlar

Bu tez çalışması kapsamında uluslararası indeksli dergilerde beş adet makale yayınlanmış [90–94] ve ayrıca bir adet yurtiçi uluslararası konferans bildirisi sunulmuştur [89].

KAYNAKLAR

[1] **Andenberg MR.**, 1973. Cluster Analysis for Applications.

[2] **Zadeh L.**, 1965. Fuzzy Sets. *Information and Control*, **8**(3), 338–353.

[3] **Ruspini EH.**, 1969. A new approach to clustering. *Information and Control*, **15**(1), 22–32.

[4] **Dunn JC.**, 1973. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, **3**(3), 32–57.

[5] **Bezdek JC.**, 2013. Pattern recognition with fuzzy objective function algorithms. Springer Science & Business Media.

[6] **Gustafson D., and Kessel W.**, 1978. Fuzzy clustering with a fuzzy covariance matrix. In *1978 IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes*, pp.761–766.

[7] **Sen S., and Dave RN.**, 1998. Clustering of relational data containing noise and outliers. In *IEEE International Conference on Fuzzy Systems*, pp.1411–1416.

[8] **Krishnapuram R., and Keller JM.**, 1993. A Possibilistic Approach to Clustering. *IEEE Transactions on Fuzzy Systems*, **1**(2), 98–110.

[9] **Pal NR., Pal K., and Bezdek JC.**, 1997. A mixed c-means clustering model. *Proceedings of 6th International Fuzzy Systems Conference*, **1**, 11–21.

[10] **Ménard M., Demko C., and Loonis P.**, 2000. The fuzzy c+2-means: solving the ambiguity rejection in clustering. *Pattern Recognition*, **33**(7), 1219–1237.

[11] **Masson MH., and Denœux T.**, 2008. ECM: An evidential version of the fuzzy c-means algorithm. *Pattern Recognition*, **41**(4), 1384–1397.

[12] **Masson MH., and Denœux T.**, 2009. RECM: Relational evidential c-means algorithm. *Pattern Recognition Letters*, **30**(11), 1015–1026.

[13] **Guo Y., and Sengur A.**, 2015. NCM: Neutrosophic c-means clustering algorithm. *Pattern Recognition*, **48**(8), 2710–2724.

[14] **Guo Y., and Sengur A.**, 2015. NECM: Neutrosophic evidential c-means clustering algorithm. *Neural Computing and Applications*, **26**(3), 561–571.

[15] **Zhang DQ., and Chen SC.**, 2003. Clustering incomplete data using kernel-based fuzzy C-means algorithm. *Neural Processing Letters*, **18**(3), 155.

[16] **Huang G-B., Zhu Q-Y., and Siew C-K.**, 2006. Extreme learning machine: Theory and applications. *Neurocomputing*, **70**(1), 489–501.

[17] **Miche Y., van Heeswijk M., Bas P., Simula O., and Lendasse A.**, 2011. TROP-ELM: A double-regularized ELM using LARS and Tikhonov regularization. *Neurocomputing*, **74**(16), 2413–2421.

[18] **Deng W., Zheng Q., and Chen L.**, 2009. Regularized Extreme Learning Machine. In *2009 IEEE Symposium on Computational Intelligence and Data Mining*, pp.389–395.

[19] **Martínez-Martínez JM., Escandell-Montero P., Soria-Olivas E., Martín-Guerrero JD., Magdalena-Benedito R., and Gómez-Sanchis J.**, 2011. Regularized extreme learning machine for regression problems. *Neurocomputing*, **74**(17), 3716–3721.

[20] **Miche Y., Sorjamaa A., Bas P., Simula O., Jutten C., and Lendasse A.**, 2010. OP-ELM: Optimally pruned extreme learning machine. *IEEE Transactions on Neural Networks*, **21**(1), 158–162.

[21] **Vajda S., and Fink GA.**, 2010. Strategies for training robust neural network based digit recognizers on unbalanced data sets. In *Proceedings - 12th International Conference on Frontiers in Handwriting Recognition, ICFHR 2010*, pp.148–153.

[22] **Mirza B., Lin Z., and Toh KA.**, 2013. Weighted online sequential extreme learning machine for class imbalance learning. *Neural Processing Letters*, **38**(3), 465–486.

[23] **Beyan C., and Fisher R.**, 2015. Classifying imbalanced data sets using similarity based hierarchical decomposition. *Pattern Recognition*, **48**(5), 1653–1672.

[24] **Chawla N V., Bowyer KW., Hall LO., and Kegelmeyer WP.**, 2002. SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, **16**, 321–357.

[25] **Pazzani M., Merz C., Murphy P., Ali K., Hume T., and Brunk C.**, 1994. Reducing Misclassification Costs. *Icml*, 217–225.

[26] **Japkowicz N.**, 2000. The Class Imbalance Problem: Significance and Strategies. *Proceedings of the 2000 International Conference on Artificial Intelligence*, 111--117.

[27] **Galar M., Fernandez A., Barrenechea E., Bustince H., and Herrera F.**, 2012. A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, **42**(4), 463–484.

[28] **He H., and Garcia EA.**, 2009. Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, **21**(9), 1263–1284.

[29] **Hu S., Liang Y., Ma L., and He Y.**, 2009. MSMOTE: Improving classification performance when training data is imbalanced. In *2nd International Workshop on Computer Science and Engineering, WCSE 2009*, pp.13–17.

[30] **Chawla N V., Lazarevic A., Hall LO., and Bowyer KW.**, 2003. SMOTEBoost: Improving Prediction of the Minority Class in Boosting. In *Knowledge Discovery in Databases: PKDD 2003: 7th European Conference on Principles and Practice of Knowledge Discovery in Databases*, Cavtat-Dubrovnik, Croatia. pp.107–119.

[31] **Barua S., Islam MM., Yao X., and Murase K.**, 2014. MWMOTE - Majority weighted minority oversampling technique for imbalanced data set learning. *IEEE Transactions on Knowledge and Data Engineering*, **26**(2), 405–425.

[32] **Radivojac P., Chawla N V., Dunker AK., and Obradovic Z.**, 2004. Classification and knowledge discovery in protein databases. *Journal of Biomedical Informatics*, **37**(4), 224–239.

[33] **Liu X-Y., Wu J., and Zhou Z-H.**, 2009. Exploratory Undersampling for Class Imbalance Learning. *IEEE Transactions on Systems, Man and Cybernetics*, **39**(2), 539–550.

[34] **Seiffert C., Khoshgoftaar TM., Van Hulse J., and Napolitano A.**, 2010. RUSBoost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, **40**(1), 185–197.

[35] **Wang BX., and Japkowicz N.**, 2010. Boosting support vector machines for imbalanced data sets. *Knowledge and Information Systems*, **25**(1), 1–20.

[36] **Tan S.**, 2005. Neighbor-weighted K-nearest neighbor for unbalanced text corpus. *Expert Systems with Applications*, **28**(4), 667–671.

[37] **Fumera G., Roli F., and Giorgio Fumera FR.**, 2002. Cost-sensitive learning in support vector machines. *VIII Convegno Associazione Italiana per L' ...*, (1).

[38] **Drummond C., Holte RC., Chawla N V., Sheng VS., Gu B., Fang W., et al.**, 2003. Exploiting the cost (in)sensitivity of decision tree splitting criteria. *International Conference on Machine Learning*, **66**(1), 239–246.

[39] **Williams DP., Myers V., and Silvious MS.**, 2009. Mine classification with imbalanced data. *IEEE Geoscience and Remote Sensing Letters*, **6**(3), 528–532.

[40] **Zong W., Huang G Bin., and Chen Y.**, 2013. Weighted extreme learning machine

for imbalance learning. *Neurocomputing*, **101**, 229–242.

[41] **Czarnecki WM.**, 2015. Weighted Tanimoto Extreme Learning Machine with Case Study in Drug Discovery. *IEEE Computational Intelligence Magazine*, **10**(3), 19–29.

[42] **Man Z., Lee K., Wang D., Cao Z., and Khoo S.**, 2013. An optimal weight learning machine for handwritten digit image recognition. *Signal Processing*, **93**(6), 1624–1638.

[43] **Mirza B., Lin Z., and Liu N.**, 2015. Ensemble of subset online sequential extreme learning machine for class imbalance and concept drift. *Neurocomputing*, (Part A), 316–329.

[44] **Zhang Y., Liu B., Cai J., and Zhang S.**, 2016. Ensemble weighted extreme learning machine for imbalanced data classification based on differential evolution. *Neural Computing and Applications*, 1–9.

[45] **Wang S., Minku LL., and Yao X.**, 2015. Resampling-based ensemble methods for online class imbalance learning. *IEEE Transactions on Knowledge and Data Engineering*, **27**(5), 1356–1368.

[46] **Fix E., and Hodges JL.**, 1951. Discriminatory Analysis - Nonparametric discrimination consistency properties. *Usaf school of aviation medicine Randolph Field, Texas*, 238–247.

[47] **Duda RO., and Hart PE.**, 1973. Pattern Classification and Scene Analysis.

[48] **Cabello D., Barro S., Salceda JM., Ruiz R., and Mira J.**, 1991. Fuzzy K-nearest neighbor classifiers for ventricular arrhythmia detection. *International Journal of Bio-Medical Computing*, **27**(2), 77–93.

[49] **Chen H-L., Yang B., Wang G., Liu J., Xu X., Wang S-J., et al.**, 2011. A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method. *Knowledge-Based Systems*, **24**(8), 1348–1359.

[50] **Chikh MA., Saidi M., and Settouti N.**, 2012. Diagnosis of Diabetes Diseases Using An Artificial Immune Recognition System2 (AIRS2) with Fuzzy K-nearest neighbor. *Journal of Medical Systems*, **36**(5), 2721–2729.

[51] **Aslan M., Akbulut Y., Şengür A., and İnce MC.**, 2017. Eklem Tabanlı Etkili Düşme Tespiti. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, **32**(4). Retrieved December 19, 2017, from http://www.mmfdergi.gazi.edu.tr/article/view/5000179780

[52] **Baoli L., Qin L., and Shiwen Y.**, 2004. An adaptive k -nearest neighbor text categorization strategy. *ACM Transactions on Asian Language Information Processing*, **3**(4), 215–226.

[53] **Keller JM., and Gray MR.**, 1985. A Fuzzy K-Nearest Neighbor Algorithm. *IEEE Transactions on Systems, Man and Cybernetics*, **SMC**-**15**(4), 580–585.

[54] **Pham TD.**, 2005. An optimally weighted fuzzy k-NN algorithm. In *Proceedings of the Third international conference on Advances in Pattern Recognition-Volume Part I*, Springer-Verlag. pp.239–247.

[55] **Denoeux T.**, 1995. A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE Transactions on Systems, Man, and Cybernetics*, **25**(5), 804–813.

[56] **Zouhal LM., and Denoeux T.**, 1998. An evidence-theoretic k-NN rule with parameter optimization. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, **28**(2), 263–271.

[57] **Zouhal LM., and Denoeux T.**, 1997. Generalizing the evidence theoretic k-NN rule to fuzzy pattern recognition. In *Proceedings of the 2nd International ICSC Symposium on Fuzzy Logic and Applications*, Zurich, Switzerland. pp.294–300.

[58] **Liu Z., Pan Q., Dezert J., Mercier G., and Liu Y.**, 2014. Fuzzy-belief K-nearest neighbor classifier for uncertain data. *Information Fusion (FUSION), 2014 17th International Conference on*, 1–8.

[59] **Liu Z-G., Pan Q., and Dezert J.**, 2013. A new belief-based K-nearest neighbor classification method. *Pattern Recognition*, **46**(3), 834–844.

[60] **Derrac J., Chiclana F., García S., and Herrera F.**, 2016. Evolutionary fuzzy k-nearest neighbors algorithm using interval-valued fuzzy sets. *Information Sciences*, **329**, 144–163.

[61] **Dudani SA.**, 1976. The Distance-Weighted k-Nearest-Neighbor Rule. *IEEE Transactions on Systems, Man and Cybernetics*, **SMC**-**6**(4), 325–327.

[62] **Gou J., Du L., Zhang Y., and Xiong T.**, 2012. A New Distance-weighted k -nearest Neighbor Classifier. *Journal of Information & Computational Science*, **9**(6), 1429–1436.

[63] **Pal NR., and Pal SK.**, 1993. A review on image segmentation techniques. *Pattern Recognition*, **26**(9), 1277–1294.

[64] **Gonzalez R., and Woods R.**, 2002. Digital image processing.

[65] **Wu Z., and Leahy R.**, 1993. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **15**(11), 1101–1113.

[66] **Smarandache F.**, 1999. A Unifying Field in Logics: Neutrosophic Logic.

[67] **Smarandache F.**, 2013. Introduction to Neutrosophic Measure, Neutrosophic Integral, and Neutrosophic Probability.

[68] **Guo Y., and Cheng HD.**, 2009. New neutrosophic approach to image segmentation. *Pattern Recognition*, **42**(5), 587–595.

[69] **Akhtar N., Agarwal N., and Burjwal A.**, 2014. K-mean algorithm for Image Segmentation using Neutrosophy. In *Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on*, pp.2417–2421.

[70] **Cheng H., Guo Y., and Zhang Y.**, 2011. A novel image segmentation approach based on neutrosophic set and improved fuzzy c-means algorithm. *New Mathematics and Natural Computation*, **7**(01), 155–171.

[71] **Zhang M., Zhang L., and Cheng HD.**, 2010. A neutrosophic approach to image segmentation based on watershed method. *Signal Processing*, **90**(5), 1510–1517.

[72] **Hanbay K., and Talu MF.**, 2014. Segmentation of SAR images using improved artificial bee colony algorithm and neutrosophic set. *Applied Soft Computing*, **21**, 433–443.

[73] **Karabatak E., Guo Y., and Sengur A.**, 2013. Modified neutrosophic approach to color image segmentation. *Journal of Electronic Imaging*, **22**(1), 013005–013005.

[74] **Sengur A., and Guo Y.**, 2011. Color texture image segmentation based on neutrosophic set and wavelet transformation. *Computer Vision and Image Understanding*, **115**(8), 1134–1144.

[75] **Mathew JM., and Simon P.**, 2014. Color texture image segmentation based on neutrosophic set and nonsubsampled contourlet transformation. In *International Conference on Applied Algorithms*, pp.164–173.

[76] **Yu B., Niu Z., and Wang L.**, 2013. Mean shift based clustering of neutrosophic domain for unsupervised constructions detection. *Optik*, **124**(21), 4697–4706.

[77] **Guo Y., and Şengür A.**, 2014. A novel image segmentation algorithm based on neutrosophic similarity clustering. *Applied Soft Computing*, **25**, 391–398.

[78] **Guo Y., Şengür A., and Ye J.**, 2014. A novel image thresholding algorithm based on neutrosophic similarity score. *Measurement*, **58**, 175–186.

[79] **Guo Y., and Şengür A.**, 2014. A novel image edge detection algorithm based on neutrosophic set. *Computers & Electrical Engineering*, **40**(8), 3–25.

[80] **Guo Y., and Şengür A.**, 2013. A novel image segmentation algorithm based on neutrosophic filtering and level set. *Neutrosophic Sets and Systems*, **1**, 46–49.

[81] **Peng B., Zhang L., and Zhang D.**, 2013. A survey of graph theoretical approaches to image segmentation. *Pattern Recognition*, **46**(3), 1020–1038.

[82] **Morris OJ., Lee M de J., and Constantinides AG.**, 1986. Graph theory for image analysis: an approach based on the shortest spanning tree. *Communications, Radar and Signal Processing, IEE Proceedings F*, **133**(2), 146–152.

[83] **Felzenszwalb PF., and Huttenlocher DP.**, 2004. Efficient graph-based image segmentation. *International Journal of Computer Vision*, **59**(2), 167–181.

[84] **Shi J., and Malik J.**, 2000. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(8), 888–905.

[85] **Wang S., and Siskind JM.**, 2003. Image segmentation with ratio cut. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **25**(6), 675–690.

[86] **Ding CH., He X., Zha H., Gu M., and Simon HD.**, 2001. A min-max cut algorithm for graph partitioning and data clustering. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pp.107–114.

[87] **Cox IJ., Rao SB., and Zhong Y.**, 1996. "Ratio regions": A technique for image segmentation. In *Proceedings - International Conference on Pattern Recognition*, pp.557–564.

[88] **Grady L.**, 2005. Multilabel random walker image segmentation using prior models. In *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, pp.763–771.

[89] **Akbulut Y., Şengür A., and Guo Y.**, 2016. Texture segmentation based on Gabor filters and neutrosophic graph cut. In *International conference on advanced technology & sciences (ICAT'16), Konya, Turkey*, pp.336–339.

[90] **Akbulut Y., Şengür A., Guo Y., and Polat K.**, 2017. KNCM: Kernel Neutrosophic c-Means Clustering. *Applied Soft Computing Journal*, **52**, 714–724.

[91] **Akbulut Y., Şengür A., Guo Y., and Smarandache F.**, 2017. A novel neutrosophic weighted extreme learning machine for imbalanced data set. *Symmetry*, **9**(8).

[92] **Akbulut Y., Sengur A., Guo Y., and Smarandache F.**, 2017. NS-k-NN:

Neutrosophic set-based k-nearest neighbors classifier. *Symmetry*, **9**(9).

[93] **Guo Y., Akbulut Y., Şengür A., Xia R., and Smarandache F.**, 2017. An efficient image segmentation algorithm using neutrosophic graph cut. *Symmetry*, **9**(9).

[94] **Guo Y., Sengur A., Akbulut Y., and Shipley A.**, 2018. An Effective Color Image Segmentation Approach Using Neutrosophic Adaptive Mean Shift Clustering. *Measurement*, **119**, 28–40.

[95] **Zhang M.**, 2010. Novel Approaches to Image Segmentation Based on Neutrosophic Logic. Utah State University.

[96] **Smarandache F.**, 2000. Collected Papers III. Oradea.

[97] **Atanassov KT.**, 1986. Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, **20**(1), 87–96.

[98] **Priest G., Sylvan R., Norman J., and Arruda AI (Ayda I.**, 1989. Paraconsistent logic.

[99] **Whittle B.**, 2004. Dialetheism, Logical Consequence and Hierarchy. *Analysis*, **64**(4), 318.

[100] **Loeb PA., and Wolff MPH.**, 2015. Nonstandard analysis for the working mathematician: Second edition.

[101] **Ju W.**, 2011. Novel Application of Neutrosophic Logic in Classifiers Evaluated Under Region-based Image Categorization System. Utah State University.

[102] **Wang H.**, 2005. Interval Neutrosophic Sets and Logic: Theory and Applications in Computing. Georgia State University.

[103] **Shan J.**, 2011. A Fully Automatic Segmentation Method For Breast Ultrasound Images. Utah State University.

[104] **Uw S., Ng AY., Jordan MI., and Weiss Y.**, 2001. On spectral clustering: Analysis and an algorithm. *Advances in Neural Information Processing Systems 14*, 849–856.

[105] **Wang Y., Jiang Y., Wu Y., and Zhou ZH.**, 2011. Spectral clustering on multiple manifolds. *IEEE Transactions on Neural Networks*, **22**(7), 1149–1161.

[106] **Yang MS., and Tsai HS.**, 2008. A Gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction. *Pattern Recognition Letters*, **29**(12), 1713–1725.

[107] **Guo Y., and Sengur A.**, 2015. A novel 3D skeleton algorithm based on neutrosophic cost function. *Applied Soft Computing Journal*, **36**, 210–217.

[108] **Smarandache F., Dezert J., Bhattacharya S., Buller A., Khoshnevisan M., Singh S., et al.**, 2001. Proceedings of the First International Conference on

Neutrosophy, Neutrosophic Logic, Neutrosophic Set, Neutrosophic Probability and Statistics. *First International Conference on Neutrosophy, Neutrosophic Logic, Set, Probability and Statistics*, (December).

[109] **Ng WWY., Hu J., Yeung DS., Yin S., and Roli F.**, 2015. Diversified sensitivity-based undersampling for imbalance classification problems. *IEEE Transactions on Cybernetics*, **45**(11), 2402–2412.

[110] **Alcalá-Fdez J., Fernández A., Luengo J., Derrac J., García S., Sánchez L., et al.**, 2011. KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic and Soft Computing*, **17**(2–3), 255–287.

[111] **Huang J., and Ling CX.**, 2005. Using AUC and accuracy in evaluating learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, **17**(3), 299–310.

[112] **Demšar J.**, 2006. Statistical Comparisons of Classifiers over Multiple Data Sets. *Journal of Machine Learning Research*, **7**, 1–30.

[113] **Hodges JL., and Lehmann EL.**, 1962. Rank Methods for Combination of Independent Experiments in Analysis of Variance. *The Annals of Mathematical Statistics*, **33**(2), 482–497.

[114] **Rodriguez-Fdez I., Canosa A., Mucientes M., and Bugarin A.**, 2015. STAC: A web platform for the comparison of algorithms using statistical tests. In *IEEE International Conference on Fuzzy Systems*.

[115] **Guo Y., and Sengur A.**, 2013. A novel color image segmentation approach based on neutrosophic set and modified fuzzy c-means. *Circuits, Systems, and Signal Processing*, **32**(4), 1699–1723.

[116] **Bache K., and Lichman M.**, 2013. UCI Machine Learning Repository. *University of California Irvine School of Information*, **2008**(14/8), 0.

[117] **MATLAB.**, 2014. R2014b. Natick, Massachusetts: The MathWorks Inc.

[118] **Yuan J., Bae E., Tai XC., and Boykov Y.**, 2010. A continuous max-flow approach to Potts model. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp.379–392.

[119] **Yasnoff WA., Mui JK., and Bacus JW.**, 1977. Error measures for scene segmentation. *Pattern Recognition*, **9**(4), 217–231.

[120] **Pratt WK.**, 2006. Digital Image Processing: PIKS Scientific Inside.

[121] **Wang S., Chung F lai., and Xiong F.**, 2008. A novel image thresholding method based on Parzen window estimate. *Pattern Recognition*, **41**(1), 117–129.

[122] **Salah M Ben., Mitiche A., and Ayed I Ben.**, 2011. Multiregion image segmentation by parametric kernel graph cuts. *IEEE Transactions on Image Processing*, **20**(2), 545–557.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER	
Adı Soyadı	Yaman AKBULUT
Doğum Yeri ve Tarihi	Elazığ – 1978
Adres	Fırat Üniversitesi Enformatik Bölümü 23119 Elazığ
E-posta	yamanakbulut@firat.edu.tr
Web	https://abs.firat.edu.tr/yamanakbulut

EĞİTİM BİLGİLERİ	
Yüksek Lisans	Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalı, 2014
Lisans	Fırat Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, 1999
Lise	Elazığ Anadolu Lisesi, 1995

MESLEKİ DENEYİMLER	
Üniversite	Fırat Üniversitesi, Enformatik Bölümü, Okutman / Öğretim Görevlisi, 2001-Devam ediyor
Özel Şirket	İvme Otomasyon, Ankara, Mühendis, 2000-2001
	Nel Elektronik, Ankara, Mühendis, 1999-2000