

AFET NEDENİYLE ZARAR GÖREN ALANLAR İÇİN YENİDEN YAPILANDIRILMIŞ BAĞIMSIZ BİLEŞENLER ANALİZİ VE KMEANS++ TEMELLİ DEĞİŞİM SAPTAMA YAKLAŞIMI

Ü. H. Atasever^a

^aErciyes Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği Bölümü, Kayseri, Türkiye, uhatasever@erciyes.edu.tr

ANAHTAR KELİMELER: Değişim Saptama, Yeniden Yapılandırılmış Bağımsız Bileşenler Analizi (RICA), Kmeans++

ÖZET:

Uzaktan algılama verileri birçok yer biliminin temel veri kaynaklarıdır. Uydu görüntüleri özellikle afetler neticesinde zarar gören alanların belirlenmesinde oldukça hızlı bulgular elde edilmesini sağlamaktadır. Bu çalışmada, optik görüntüler için afet nedeniyle zarar gören alanları çok hızlı biçimde hesaplayabilen kontrolsüz bir değişim saptama yaklaşımı önerilmiştir. Bu çalışmanın en önemli özellikleri çok hızlı ve etkin olmasına ek olarak çok az kontrol parametresine ihtiyaç duymasındır. Başlangıç olarak afet öncesi ve sonrası görüntüleri ile fark görüntüleri hesaplanır. Ardından bu fark görüntülerine izole piksellerin giderilebilmesi için wiener filtre uygulanır. Daha sonra filtrelenmiş fark görüntüleri Yeniden Yapılandırılmış Bağımsız Bileşenler Analizi ile tek boyutlu bir öznitelik uzayına projekte edilir. Son aşamada ise tek boyutlu veri Kmeans++ algoritmasıyla kümelenecek, değişen ve değişmeyen bölgeler belirlenmiş olur. Sunulan yaklaşım Sardinia ve Mexico veri setleri kullanılarak test edilmiştir. Literatürde kabul görmüş farklı kontrolsüz değişim saptama teknikleri ile kıyaslanarak etkinliği gösterilmiştir.

UNSUPERVISED CHANGE DETECTION APPROACH BASED ON RECONSTRUCTION INDEPENDENT COMPONENT ANALYSIS AND KMEANS++ FOR DAMAGED AREAS BY DISASTER

KEYWORDS: Change Detection, Reconstruction Independent Component Analysis (RICA), Kmeans++

ABSTRACT:

Remote sensing data is one of the basic data sources of many earth sciences. Satellite images provide very rapid findings in determining the areas that are damaged due to disasters. In this study, an unsupervised change detection approach is proposed which can calculate the areas affected by the disaster very quickly for optical images. The most important features of this study are very fast, effective and require very few control parameters. Initially, difference images are calculated with pre-disaster and post-disaster images. Then, wiener filter is applied to this difference images, separately to remove the isolated pixels. The filtered difference images are then projected into a one-dimensional attribute space with Reconstruction Independent Component Analysis. In the final stage, one-dimensional data is clustered with the Kmeans ++ algorithm and the changed and unchanged regions are determined. The presented approach has been tested using Sardinia and Mexico data sets. Effectiveness of proposed approach has been shown by comprising many unsupervised change detection approach.

1. GİRİŞ

Uzaktan algılama düzenli veri temini sağlayan, çok geniş alanları gözleme kabiliyetine haiz ve bu verilerden anlamlı tematik bilgilerin çıkarılabilmesini mümkün kılan bir bilim dalıdır. Bu özellikleri nedeniyle, tarım, çevre, jeoloji, jeofizik, ormancılık gibi farklı bilim dallarının temel veri kaynaklarından biri haline gelmiştir (Ghosh *et al.* 2011, Gong *et al.* 2012, Subudhi *et al.* 2014). Arazi örtüsünde meydana gelen değişimlerin saptanması uzaktan algılama biliminin en önemli araştırma konularından biridir. Kentsel büyüme, afet yönetimi, çölleşme ve bitki gelişimlerinin izlenmesi gibi uygulamalarda uydu görüntülerinin kullanımı oldukça önemlidir (Celik 2009). Özellikle bir afet yaşanan bölgelerde zarar gören alanların tespiti için değişim saptama tekniklerinin kullanımı sıkça başvurulan bir çözümdür. Afet öncesi ve afet sonrası görüntüleri kullanılarak gerçekleştirilen bu uygulamalar zamanında ve hızlı kararlar alabilmek için oldukça önemli bir rol oynamaktadır. Genellikle kontrollü ve kontrolsüz olarak ikiye ayrılan değişim

saptama tekniklerinin birbirlerine göre avantaj ve dezavantajları bulunmaktadır. Kontrollü değişim saptama tekniklerinde çoğunlukla kontrollü sınıflandırma metodları kullanıldığı için her bir sınıfın ne kadar değiştiğini belirlemek görece daha kolaydır. Ancak bu uygulamalarda değişim saptama işleminin başarısı sınıflandırma tekniğinin parametrelerinin doğru biçimde ayarlanmasına oldukça bağlıdır. Destek vektör makineleri, yapay sinir ağları, derin öğrenme ve hızlandırma teknikleri oldukça başarılı sınıflandırma yapabilmektedir. Ancak bu başarılı ve etkin yöntemlerin ayarları tecrübe ve uzmanlık gerektirmesi nedeniyle farklı disiplinlerdeki insanların kullanımını oldukça güçleştirmektedir. Öte yandan, kontrolsüz değişim saptama tekniklerinde sınıf değişimleri belirlemek zor olmasına karşın parametre sayısı genellikle azdır. Üstelik eğitim için alan seçimi gerekmemesi de süreci kolaylaştırmaktadır. Kontrolsüz değişim saptama teknikleri genellikle özellik uzayını çıkartımı ve kümeleme/segmentasyon aşaması olmak üzere iki ana adımdan oluşur (Neagoe *et al.* 2014). Bu iki temel

adım içinde uygulanacak veya kullanılacak teknikler model doğruluğunu etkileyen en önemli unsurlardandır. Önerilen yaklaşım özellik uzayı çıkartımın ciddi bir yenilik içermektedir. Afet öncesi ve afet sonrası görüntülerden standart fark görüntüsü ve logaritmik fark görüntüleri bir öznitelik çıkartım tekniği olan Yeniden Yapılandırılmış Bağımsız Bileşenler Analizi ile tek bantlı bir veri haline getirilmiş ve ardından Kmeans++ ile kümelenecek değişen ve değişmeyen alanlar bulunmuştur. Bu öznitelik çıkartım tekniği sayesinde iki fark görüntüsünden çok ciddi bilgi kaybı olmadan veri boyutu yarıya indirilmiş ve Kmeans++ algoritmasının yakınsaması daha hızlı hale getirilmiştir. Kmeans++ temelde standart Kmeans algoritmasının küme merkezi başlangıç değerleri için yeni bir strateji içermektedir (Arthur *et al.* 2007). Ayrıca yapılan çalışmalarda standart Kmeans'e göre daha hızlı yakınsadığı ve daha başarılı kümeleme yapabildiği gösterilmiştir (Arthur and Vassilvitskii 2007). Önerilen yaklaşımın etkinliğini incelemek için farklı afetlerden sonra üretilmiş Sardinia ve Mexico veri setleri kullanılmıştır.

2. ÖNERİLEN YAKLAŞIM

Önerilen yaklaşım tek bant ile çalışmak üzere hazırlanmıştır. Bu nedenle değişimin belirlenebileceği uygun band seçilmelidir. Ardında afet öncesi ve sonrası görüntüleri birbirine kaydedilir. Eğer gerekliyse radyometrik düzeltme yapılmalıdır. Önerilen yöntemin ilk aşamasında iki farklı fonksiyonla fark görüntüleri oluşturulur:

$$img_{F1} = |img_2 - img_1| \quad (1)$$

$$img_{F2} = \left| \log \left(\frac{img_2 + 1}{img_1 + 1} \right) \right| \quad (2)$$

Eşitliklerde img_{F1} standart fark görüntüsü, img_{F2} logaritmik oran/fark görüntüsü, img_1 afet öncesi görüntü ve img_2 afet sonrası görüntüdür. Ardından fark görüntülerine 2 boyutlu wiener filtre uygulanır. Burada önemli olan uygun pencere büyüklüğünü belirleyebilmektedir. Farklı görüntüler için farklı pencere büyüklükleri kullanmak gerekebilmektedir. Ancak bu çalışmada [23 23] pencere büyüklüğünün iki veri seti için de oldukça iyi sonuçlar üretebildiği gözlemlenmiştir.

Filtreleme işleminin ardından fark görüntüleri öznitelik çıkartım tekniklerinden olan Yeniden Yapılandırılmış Bağımsız Bileşenler Analizi ile bir sütun vektör üretilir.

Yeniden Yapılandırılmış Bağımsız Bileşenler Analizi (YYBBA) bir yeniden yapılandırma fonksiyonunun minimizasyonunu sağlayarak çözüm üretmektedir. Standart ICA model Eşitlik 3'de sunulmuştur (Le *et al.* 2011, Lei *et al.* 2016).

$$x = \beta + A \cdot s \quad (3)$$

Eşitlikte x , k uzunluğundaki bir sütun vektör; β k uzunluğunda ve sabit bir terimi temsil eden bir sütun vektör; s uzunluğu q olan ve sıfır ortalamalı olan bir sütun vektörü; A ise $k \times q$ boyutunda karışım matrisini göstermektedir. Amaç bilinmeyen A ve s değerlerinin gözlem vektörü x kullanılarak kestirilmesidir. Eşitlik 1 ve 2 ile hesaplanmış fark görüntüleri $r \times c$ büyüklüğünde sütun matrisine dönüştürülür ve temel veri matrisi Eşitlik 4'deki gibi oluşturulur (Do *et al.* 2007).

$$X = \begin{bmatrix} x_1^1 & x_1^2 \\ x_2^1 & x_2^2 \\ x_3^1 & x_3^2 \\ \vdots & \vdots \\ x_n^1 & x_n^2 \end{bmatrix} \quad (4)$$

Eşitlik 4'deki her bir sütun bir özniteligi yani fark görüntüsünü temsil etmektedir. Ardından model her bir öznitelik için bir adet w ağırlık katsayısı hesaplar Eşitlik 5'de verilmiş olan ağırlık matrisi oluşur.

$$W = [w_1 \quad w_2] \quad (5)$$

YYBBA, standart ICA'dan farklı olarak Quasi Newton optimizasyon tekniklerinden Broyden-Fletcher-Goldfarb-Shanno (BFGS) algoritmasını kullanarak RICA maliyet fonksiyonunu minimize eder. Maliyet fonksiyonu Eşitlik 6'da verilmiştir (Le, Karpenko, Ngiam and Ng 2011).

$$h = \frac{\lambda}{n} \sum_{i=1}^n \|WW^T x_i - x_i\|_2^2 + \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^q \sigma_j g(w_j^T x_i) \quad (6)$$

Eşitlik 6'da λ regülasyon parametresidir ve bu çalışmada 1 alınmıştır. σ standartlaşma sürecindeki standart sapma değerleridir. $g()$ fonksiyonu kontrast fonksiyonudur. Kontrast fonksiyonu için farklı seçenekler olmasına rağmen bu çalışmada Eşitlik 7'de verilen kullanılmıştır (Le, Karpenko, Ngiam and Ng 2011).

$$g = \frac{1}{2} \log(\cosh(2x)) \quad (7)$$

Ağırlıkların bulunmasının ardından 2-D öznitelik verisi 1-D'ye dönüştürülür. Son adımda ise elde edilen 1-D veri Kmeans++ algoritması ile değişen ve değişmeyen alanlar olmak üzere iki sınıfa kümelendir. Kmeans++ büyük oranda standart Kmeans'e benzerine rağmen küme merkezleri için kullanılan başlangıç değerlerinin seçiminde farklı bir strateji izlemektedir. Kmeans++ küme merkezlerini bulmak için sezgisel bir tarama kullanır. Yapılan çalışmalarda standart Kmeans'e göre hem süre bakımından hem çözümün kalitesi bakımında daha iyi çözümler ürettiği gözlemlenmiştir. Sınıf sayısı k olduğu varsayılan bir Kmeans++ işlem adımları şu şekildedir (Arthur and Vassilvitskii 2007):

1. Veri kümesinden rastgele bir gözlem (c_1) seçilir.
2. Her bir gözlem değerinin c_1 'e olan uzaklığı hesaplanır. c_j ve m gözlemi arasındaki uzaklık $dist(x_m, c_j)$ olsun.
3. Diğer sınıf merkezi, c_2 , Eşitlik 8' de hesaplanan olasılık değerine göre seçilir:

$$\frac{dist^2(x_m, c_1)}{\sum_{j=1}^n dist^2(x_j, c_1)} \quad (8)$$

4. Küme merkezini seçmek için
 - a. Her bir gözlemin her bir sınıf merkezine olan uzaklığı hesaplanır ve gözlemler en yakın oldukları sınıflara atanırlar

- b. $m = 1, \dots, n$ ve $p = 1, \dots, j-1$ olmak üzere kümeye merkezi j Eşitlik 9'da hesaplanan olasılığa göre seçilir

$$\frac{\text{dist}^2(x_m, c_p)}{\sum_{\{h: x_h \in C_p\}} \text{dist}^2(x_h, c_p)} \quad (9)$$

Eşitlikte C_p küme merkezi c_p 'ye en yakın gözlemlerin tamamıdır.

5. k küme merkezleri seçilene kadar 4. Adım tekrar edilir.

3. DENEYSSEL ÇALIŞMALAR

Önerilen yaklaşımın performansını incelemek için daha önce benzer çalışmalarda da kıyas amacıyla kullanılmış olan optik Sardinia ve Mexico veri setleri kullanılmıştır (Atasever *et al.* 2016, Atasever 2017, Ghosh, Mishra and Ghosh 2011, Hao *et al.* 2013, Mishra *et al.* 2012). İki veri setindeki görüntüler birbirine kaydedilmiştir. Ayrıca histogramlarının düzgün olması nedeniyle radyometrik düzeltme yapılmasına gerek duyulmamıştır. Bu iki veri setinde, çalışma bölgelerine ait afet öncesi, afet sonrası ve yer gerçeği görüntüsü olmak üzere toplam üç görüntü bulunmaktadır. Yer gerçeği görüntüsü yaşanan afet nedeniyle zarar gören alanın tam tahribatını içerecek biçimde hazırlanmıştır. Yapılan testler için görüntüleri ait yakın kızılötesi bantlar tercih edilmiştir. Deneysel uygulamada önerilen yaklaşım HCM, PCA-DS, PCA K-means, PCA-FCM, ile kıyaslanmıştır. Karşılaştırma kriteri olarak Yanlış Alarm (False Alarm), Eksik Alarm (Missed Alarm), Toplam Hata ve Toplam Hata Oranı kullanılmıştır. Bu kriterler birçok farklı çalışmada temel değerlendirme kriteri olarak kullanılmıştır. Bu kriterler ise şu anlamlara gelmektedirler:

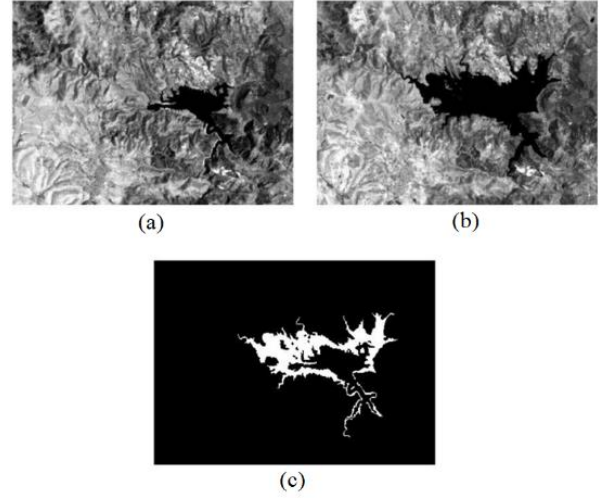
- Eksik (Missed) Alarm: Değişmemiş olarak etiketlenmiş değişen piksel sayısı
- Yanlış (False) Alarm: Değişmiş olarak etiketlenmiş değişmeyen piksel sayısı
- Toplam Hata: Yanlış etiketlenmiş piksellerin sayısı, Eksik Alarm ve Yanlış Alarm Toplamı
- Toplam Hata Oranı: Toplam hata değerinin görüntüdeki tüm piksel sayısına oranının yüz katı.

3.1 Sardinia Veri Seti ile Uygulama

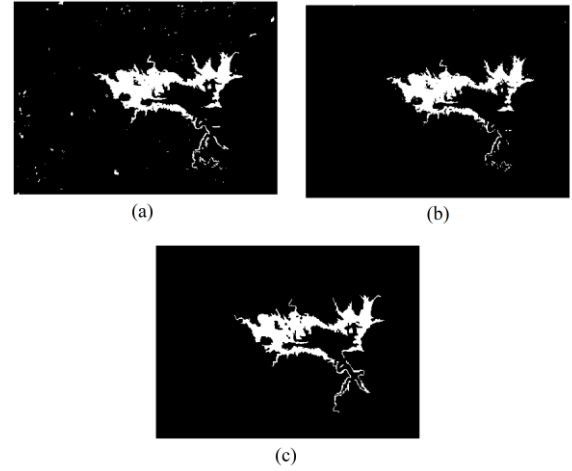
İtalya'nın Sardinia Adası'ndaki Mulargia gölüne ait Landsat 5(TM) görüntülerinden oluşan Sardinia veri seti 300x412 çözünürlüğe sahiptir. Sardinia veri setinde 1995, 1996 yıllarına ait görüntülerle birlikte yer gerçeği görüntüsü bulunmaktadır. Veri setine ilişkin görüntüler Şekil 1'de sunulmuştur. Bu veri seti Mulargia Gölü'nün taşması ve çevresinin etkilenmesi sonucu oluşan değişim nedeniyle ortaya çıkmış ve araştırmacılar tarafından hazırlanmıştır. Elde edilen kantitatif bulgular karşılaştırmalı olarak Tablo 1'de verilmiştir. Tablo incelendiğinde karşılaştırılan yöntemler içinde en iyi toplam hata değeri ve toplam hata oranı değerlerini önerilen yaklaşımın elde ettiği gözlemlenmiştir. Ayrıca PCA-Kmeans yaklaşımı ile görsel olarak karşılaştırabilmek için elde edilen değişim haritası Şekil 2' de sunulmuştur. Elde edilen değişim haritaları dikkatli biçimde incelenirse, sunulan yaklaşımın yer gerçeği haritasına çok daha benzer bir olduğu görülmektedir.

3.2 Mexico Veri Seti ile Uygulama

Mexico veri seti, Meksika'da bir bölgeden alınmış Nisan 2000 ve Mayıs 2002 tarihlerine ait Landsat 7 ETM+ görüntüleri ile bir yer gerçeği görüntüsünü içermektedir. Bir orman yangını neticesinde oldukça geniş bir alandaki bitki örtüsü yok olmuştur. Veri setindeki görüntülerin boyutu ise 512x512 pikseldir. Mexico veri setine ilişkin görüntüler Şekil 3'de sunulmuştur. Tablo 3'de gösterildiği gibi, sunulan yaklaşım oldukça iyi değişim saptama doğruluğu elde etmiştir. Ek olarak önerilen yaklaşımın performansının Mexico veri seti için görsel olarak irdelenebilmesi için elde edilen değişim haritaları da Şekil 4'de verilmiştir. Bu görselde PCA-Kmeans yaklaşımı ile elde edilen değişim haritası da bulunmaktadır.



Şekil 1. Sardinia (Mulargia) Veri seti a) Eylül 1995, b) Eylül 1996, c) Yer gerçeği görüntüsü



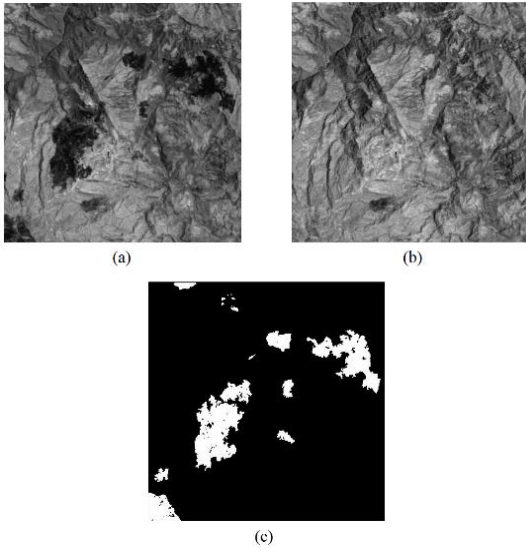
Şekil 2. Sardinia veri seti için elde edilen sonuçlar a) PCA-Kmeans, b) Önerilen Yaklaşım, c) Yer gerçeği görüntüsü

Metot	Yanlış Alarm	Eksik Alarm	Toplam Hata	Toplam Hata Oranı
HCM (Ghosh, Mishra and Ghosh 2011)	275	4133	4408	3.56
PCA-FCM	752	1855	2607	2.10
PCA-Kmeans (Celik 2009)	1603	832	2435	1.97
PCA-DS (Atasever, Kesikoglu and Ozkan 2016)	744	1405	2159	1.73
Önerilen Yaklaşım	949	1078	2027	1.64

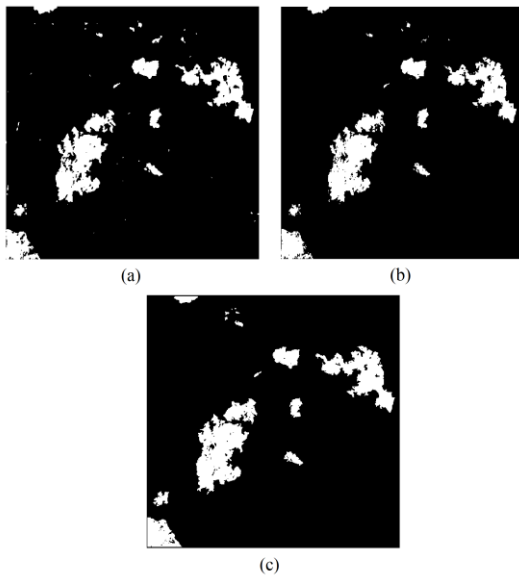
Tablo 1. Sardinia veri seti için farklı yaklaşımlara ait değişim saptama performansları

Metot	Yanlış Alarm	Eksik Alarm	Toplam Hata	Toplam Hata Oranı
PCA-FCM	3729	1008	4737	1.80
PCA-Kmeans (Celik 2009)	3662	1024	4686	1.78
HCM(Ghosh, Mishra and Ghosh 2011)	747	3425	4172	1.59
Önerilen Yaklaşım	720	3298	4018	1.53
PCA-DS(Atasever, Kesikoglu and Ozkan 2016)	2428	1589	4017	1.53

Tablo 2. Mexico veri seti için farklı yaklaşımlara ait değişim saptama performansları



Şekil 3. Mexico veri seti a) Nisan 2000 görüntüsü b) Mayıs 2002 görüntüsü c) Yer gerçeği görüntüsü



Şekil 4. Sardinia veri seti için elde edilen sonuçlar a) PCA-Kmeans, b) Önerilen Yaklaşım, c) Yer gerçeği görüntüsü

4. SONUÇLAR

Değişim saptama birçok farklı disiplinin en önemli çalışma konularından biridir. Bu çalışmada oldukça hızlı çalışabilen etkin bir kontrolsüz değişim saptama yaklaşımı önerilmiştir. Öznitelik çıkartım aşamasında Yeniden Yapılandırılmış Bağımsız Bileşenler Analizi tekniği kullanılmış ve değişen değişmeyen alanların tespiti için Kmeans++ kümeleme algoritması kullanılmıştır. Kmeans++ algoritması standart Kmeans algoritmasına göre oldukça hızlı yakınsama özelliğine sahip bir algoritmadır. Farklı veri setleri ile yapılan uygulamalar da hem kalitatif hem de kantitatif olarak önerilen yaklaşımın oldukça etkin sonuçlar üretebildiği iki farklı veri seti ile gösterilmiştir.

KAYNAKLAR

- Arthur, D. and Vassilvitskii, S., 2007. k-means++: the advantages of careful seeding, *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, Society for Industrial and Applied Mathematics, New Orleans, Louisiana, pp. 1027-1035.
- Atasever, U.H., Kesikoglu, M.H. and Ozkan, C., 2016. A New Artificial Intelligence Optimization Method for PCA Based Unsupervised Change Detection of Remote Sensing Image Data, *Neural Network World*, 26(2):141-154.
- Atasever, U.H., 2017. A New Unsupervised Change Detection Approach with Hybrid Clustering for Detecting The Areal Damage After Natural Disaster, *Fresenius Environmental Bulletin*, 26(6):3891-3896.
- Celik, T., 2009. Multiscale Change Detection in Multitemporal Satellite Images, *IEEE Geoscience and Remote Sensing Letters*, 6(4):820-824.
- Celik, T., 2009. Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and K-Means Clustering, *IEEE Geoscience and Remote Sensing Letters*, 6(4):772-776.
- Do, C.M. and Javidi, B., 2007. Multifocus Holographic 3-D Image Fusion Using Independent Component Analysis, *Journal of Display Technology*, 3(3):326-332.
- Ghosh, A., Mishra, N.S. and Ghosh, S., 2011. Fuzzy Clustering Algorithms for Unsupervised Change Detection in Remote Sensing Images, *Information Sciences*, 181(4):699-715.
- Gong, M., Zhou, Z. and Ma, J., 2012. Change Detection in Synthetic Aperture Radar Images based on Image Fusion and

- Fuzzy Clustering, *IEEE Transactions on Image Processing*, 21(4):2141-2151.
- Hao, M., Zhang, H., Shi, W. and Deng, K., 2013. Unsupervised Change Detection Using Fuzzy C-Means and MRF From Remotely Sensed Images, *Remote Sensing Letters*, 4(12):1185-1194.
- Le, Q.V., Karpenko, A., Ngiam, J. and Ng, A.Y., 2011. ICA with reconstruction cost for efficient overcomplete feature learning, *Proceedings of the 24th International Conference on Neural Information Processing Systems*, Curran Associates Inc., Granada, Spain, pp. 1017-1025.
- Lei, Y., Shan, H., Jia, F. and Lin, J. Reconstruction independent component analysis-based methods for intelligent fault diagnosis, *Proceedings of the 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 4-6 May 2016 2016, pp. 245-250.
- Mishra, N.S., Ghosh, S. and Ghosh, A., 2012. Fuzzy clustering algorithms incorporating local information for change detection in remotely sensed images, *Applied Soft Computing*, 12(8):2683-2692.
- Neagoe, V.E., Stoica, R.M., Ciurea, A.I., Bruzzone, L. and Bovolo, F., 2014. Concurrent Self-Organizing Maps for Supervised/Unsupervised Change Detection in Remote Sensing Images, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(8):3525-3533.
- Subudhi, B.N., Bovolo, F., Ghosh, A. and Bruzzone, L., 2014. Spatio-Contextual Fuzzy Clustering with Markov Random Field Model for Change Detection in Remotely Sensed Images, *Optics & Laser Technology*, 57:284-292.