



Histopatolojik Görüntülerde Hücre Bölütlemesinde Değişimsel Kip Ayrışımı ve Görgül Kip Ayrışımı Özelliklerinin Karşılaştırılması

Comparison of Variational Mode Decomposition and Empirical Mode Decomposition Features for Cell Segmentation in Histopathological Images

Omer Faruk Karaaslan, Gokhan Bilgin
Bilgisayar Mühendisliği Bölümü
Yıldız Teknik Üniversitesi
34220 İstanbul, Türkiye
{omerfk,gbilgin}@yildiz.edu.tr

Özetçe —Bu çalışmada sayısal histopatolojik görüntülerindeki hücrelerin veri uyumlu özellik çıkarım yöntemleriyle bölütlenme başarımının artırılması amaçlanmıştır. Bu amaç doğrultusunda görgül kip ayrışımı ve değişimsel kip ayrışımı yöntemlerinin karşılaştırmalı olarak kullanılması hedeflenmiştir. İlk olarak sayısal histopatolojik görüntülerin RGB renk uzayından gri düzeyeye dönüşümü gerçekleştirilmiştir. Ardından görgül kip ayrışımı ve değişimsel kip ayrışımı yöntemleri bu görüntülere uygulanmış ve elde edilen öznelikler çekirdek tabanlı sınıflandırıcı olan destek vektör makineleri ve topluluk tabanlı sınıflayıcı olan rassal ormanlar yöntemleri ile sınıflandırılmıştır. Sonuçlar üç farklı ölçüte göre değerlendirilmiştir. Uygulama sonuçları bölümünde, bu çalışmada elde edilen sonuçlar tüm ayrıntılarıyla sunulmuştur.

Anahtar Kelimeler—Hücre bölütleme, histopatolojik görüntü analizi, görgül kip ayrışımı, değişimsel kip ayrışımı.

Abstract—In this study, it is aimed to increase the segmentation performance of the cells in the digital histopathological images by data compatible feature extraction methods. For this purpose, it is proposed to use empirical mode decomposition and variational mode decomposition methods as a comparison. Initially, the conversion of digital histopathological images from RGB color space to gray level is performed. Then, empirical mode decomposition and variational mode decomposition methods are applied to these images, and the obtained features are classified by using support vector machines which is a kernel-based classifier and random forests which is an ensemble-based classifier. The results are evaluated according to three different metrics. In the application results section, the results obtained in this study are presented in detail.

Keywords—Cell segmentation, histopathological image analysis, empirical mode decomposition, variational mode decomposition.

I. GİRİŞ

Günümüzde kanser, ölümcül hastalıklar sıralamasında kalp krizinden sonra ikinci sırada yer almaktadır. Uluslararası Kanser Araştırmaları Ajansı verilerine göre dünya çapında yaklaşık 28 milyon insan, çeşitli kanser türlerine (akciğer, göğüs, prostat ve kolon) yakalanarak hayatını kaybetmiştir [1]. Ölüm oranı oldukça yüksek olan bu hastalığın erken teşhisinin, hastanın motivasyonunu güçlendirmek ve tedavi sürecinin hızlandırılması açısından önemi oldukça büyüktür.

Yunanca doku (histos), hastalık (pathos) ve çalışma (logos) kelimelerinden oluşan histopatoloji, ilgili uzmana erken teşhis konusunda yardımcı olarak geliştirilen bir tıp alanıdır [2]. Histopatolojide doku örneklerindeki kanserli hücre miktarı ve büyüme hızı Nottingham [3] ve Gleason [4] gibi belirli standartlara göre ölçülmekte, buna göre hasta gerekli tedaviye yönlendirilmektedir.

Histopatolojik görüntülerin bilgisayar destekli sistemler sayesinde ayrıştırıcı niteliklerinin ortaya çıkarılması, tedavi sürecini olumlu yönde etkilemektedir. Bu amaçla sayısal histopatoloji alanında çeşitli çalışmalar yapılmıştır [5]–[7]. Yapılan çalışmalar incelendiğinde, ağırlıklı olarak ayrıştırıcı özellik çıkarımı için Fourier ve Dalgacık Dönüşümü gibi temel işaret işleme yöntemlerinin kullanıldığı gözlenmiştir [8]. Fourier Dönüşümü (FD), sadece frekans düzleminde ve durağan işaretlerde kullanışlı olduğundan ve uzamsal bilginin önemi, kanserli yapıların karakteristiğini belirleyecek düzeyde yüksek olduğundan [9], histopatolojik görüntülerde FD kullanımını kısıtlayıcı bir etkiye sahip olmaktadır. FD'ye alternatif olarak öne sürülen Dalgacık Dönüşümü (DD), frekans bilgisinin uzamsal yapısını da ayrıştırabilmesine karşın, veriye uygun dalgacık fonksiyonu seçiminde net bir yöntem barındırmamakta, bu da kanserli hücre tespit sürecinin uzamasına sebep olmaktadır.

Yakın zamanda ortaya çıkan Hilbert-Huang Dönüşümü ve onun ilk aşaması olan Görgül Kip Ayrışımı (GKA), FD ya da DD gibi belirli bir ana fonksiyona bağlı kalmaksızın, veriyle uyumlu, yinelemeli bir şekilde frekans ve uzam bilgisinin çıkarılmasını sağlayan bir yöntemdir [10]. Gürültüye dayanıklılığının artırılması amacıyla "Topluluk Temelli Görgül Kip Ayrışımı" sürümü ortaya konulmuştur [11]. Buna karşın matematiksel bir yapısının olmaması, yinelemeyi sona erdirmeye kısıtının belirli olmaması, işlem karmaşıklığını da beraberinde getirmektedir. Ayrıca GKA için en büyük sorunlardan biri de, kip karışımı olarak bilinen olgudur [12]. Bu olguya göre birden farklı bileşen, bir İçkin Kip Fonksiyonu (IKF) içinde yer alabilmekte, bu da işaretten anlamlı özniteliklerin ayrıştırılması sürecini zorlaştırabilmektedir.

GKA'nın eksikliklerini giderebilmek adına 2014'te Değişimsel Kip Ayrışımı (DKA) sunulmuştur [13]. DKA, merkezi frekanslar üzerine ayrıştırmayı sağlayan bir yapıdadır. Ayrıca her bileşen, sadece 1 yineleme ile elde edilir, bu da işaretin ayrıştırılmasını hızlandırır. Son zamanlarda işaretlerden özellik çıkarımının [14], [15] yanında, iki boyutlu sürümüyle [16] biyomedikal görüntülerden özellik çıkarımı konusunda da kullanılmaya başlanmıştır [17], [18].

Bu çalışmanın 2. bölümünde, GKA ve DKA ana hatlarıyla anlatılmıştır. "Uygulama Sonuçları" bölümünde, kullanılacak veri kümesi sunulmuş, ardından bu veri kümesinden üç farklı özellik kümesinin Rassal Ormanlar ve Destek Vektör Makineleri kullanılarak sınıflandırma başarımları incelenmiştir. "Sonuç" bölümünde de kullanılmakta olan DKA ve GKA yöntemlerinin, sınıflandırma başarımlarını orijinal görüntü verilerine göre daha fazla arttırmayı sağlayacak özellikler sunduğu ortaya konulmuştur, ileride yapılması planlanan çalışmalar öne sürülmüştür.

II. DEĞİŞİMSSEL KIP AYRIŞIMI

DKA, 1998'de ortaya konulan GKA'nın eksik yönlerinin tamamlanması amacıyla oluşturulan bir işaret ayrıştırma yöntemidir [13]. Bu yöntemin anlaşılabilir olması için öncelikle GKA'nın adımları incelenmesi gereklidir:

- 1) Zaman değişkenine (t) bağlı ilgili işaretin ($x(t)$) tüm yerel minimum ve maksimum noktaları elde edilir. Maksimum noktalar üst zarfı (env_{max}), minimum noktalar da alt zarfı (env_{min}) oluşturur.
- 2) Elde edilen zarfların ortalaması alınır ve giriş işaretinden çıkarılır:

$$x(t) = x(t) - \frac{env_{max} + env_{min}}{2} \quad (1)$$

- 3) Ortalama zarf işareti uç noktalar barındırdıkça bu işlemler birinci adımdan itibaren devam ettirilir. Uç nokta kalmadığında İKF elde edilmiş olur.
- 4) Bulunan İKF, giriş işaretinden çıkarılır ve kalan işaret, giriş işareti gibi değerlendirilerek işlem birinci adımdan devam ettirilir. Son durumda işaret uç nokta barındırmadığı zaman, artık işareti (art) elde edilir.

Bu adımlar sonunda elde edilen yapıda, k adet İKF ve bir adet artık işareti mevcut olur:

$$x(t) = \sum_{i=1}^k IKF_i + art \quad (2)$$

DKA'da ise her kip (u_k), bir merkezi frekans etrafında (w_k) öbeklenir. DKA bu kip ve merkezi frekansları,

$$x(t) = \sum_{k=1}^K u_k(t) \quad (3)$$

koşulunu sağlamak şartıyla, aşağıda belirtilen kısıtlanmalı problemi çözerek elde eder:

$$\min_{u_k, w_k} \sum_{k=1}^K \left\| \partial_t \left[(\delta(t) + \frac{j}{\pi t}) * u_k(t) \right] e^{-jw_k t} \right\|_2^2 \quad (4)$$

Burada δ , Dirac dağılımını ve $*$ da konvolüsyon işlemi ifade eder. Lagrange çarpanı (λ) ve veri dengeleme parametresiyle (α) problem, kısıtlamasız hâle dönüşür ve alternatif yönlü çarpanlar yöntemiyle çözümü elde edilir:

$$L(u_k, w_k, \lambda) = \alpha \sum_{k=1}^K \left\| \partial_t \left[(\delta(t) + \frac{j}{\pi t}) * u_k(t) \right] e^{-jw_k t} \right\|_2^2 + \left\| x(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \langle \lambda(t), x(t) - \sum_{k=1}^K u_k(t) \rangle \quad (5)$$

Elde edilen optimum $u_k(w)$ değerleri frekans uzamında Wiener süzgeciyle güncellenir. Böylece gürültüye karşı dayanıklılık sağlanır. Kipin zaman uzamındaki karşılığına ($u_k(t)$) da şu şekilde erişilir:

$$u_k^{n+1}(w) = \frac{x_f(w) - \sum_{i \neq k} u_i(w) + \frac{\lambda(w)}{2}}{1 + 2\alpha(w - w_k)^2} \quad (6)$$

$$u_k(t) = \mathfrak{R}(ifft(u_k(w))) \quad (7)$$

$$w_k^{n+1} = \frac{\int_0^\infty w |u_k(w)|^2 dw}{\int_0^\infty |u_k(w)|^2 dw} \quad (8)$$

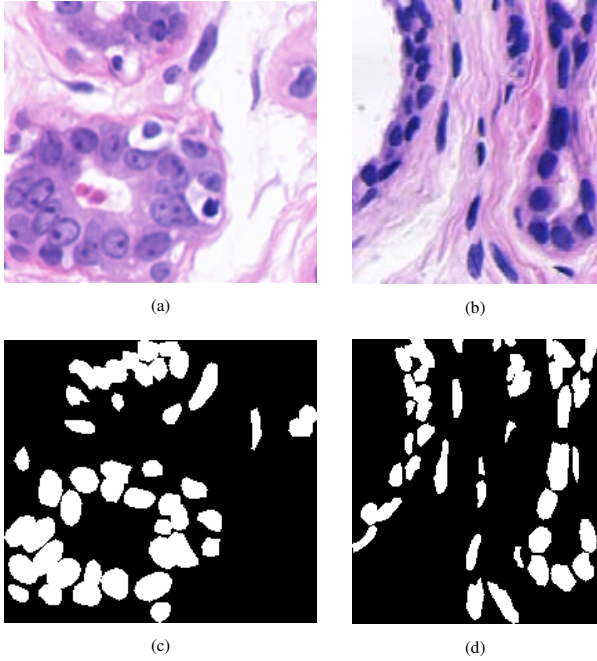
III. UYGULAMA SONUÇLARI

Bu çalışmada Yale Üniversitesi'ndeki David Rimm Laboratuvarı tarafından sağlanan histopatolojik görüntüler kullanılmıştır [19]. Bu veri kümesinde 32 iyi huylu ve 26 kötü huylu tümör görüntüsü mevcuttur. Hücre çekirdekleri Hematoksilin ile, diğer yapılar Eozin ile boyanmıştır. Görüntüler RGB renk uzayında hazırlanmıştır. Her görüntünün kendine ait, hücre piksellerinin beyaz, diğer hücre yapılarının siyah olarak gösterildiği bir yer gerçekliği (ground-truth) vardır. Veri kümesinden iki örnek görüntü, Şekil 1'de gösterilmiştir.

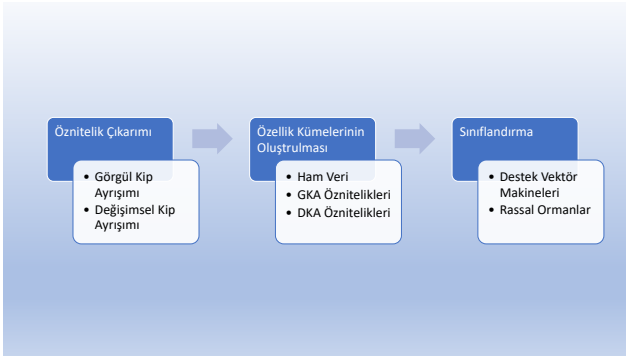
Uygulamanın genel akış şeması Şekil 2'de gösterilmiştir. İlk aşamada, 3 farklı özellik kümesi oluşturulmuştur. *Özellik Kümesi 1*'de (ÖK-1), görüntülerin gri-düzye değerleri kullanılmıştır. *Özellik Kümesi 2*'de (ÖK-2), gri-düzye görüntülere GKA uygulanmıştır. *Özellik Kümesi 3*'te (ÖK-3), gri-düzye görüntülere DKA uygulanmıştır (Tablo I).

Tablo I: Oluşturulan özellik kümelerinin içerikleri

Özellik Kümesi	Açıklama
ÖK-1	Görüntülerin gri-düzye piksel değerleri kullanılmıştır.
ÖK-2	Görüntülerin gri-düzye piksel değerlerine GKA uygulanmıştır.
ÖK-3	Görüntülerin gri-düzye piksel değerlerine DKA uygulanmıştır.



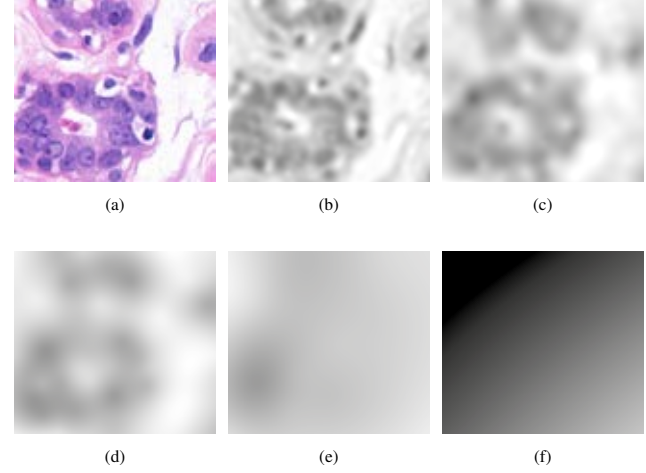
Şekil 1: Uygulamada kullanılan veri kümesine ait 2 örnek görüntü ((a) iyi huylu tümör örneği, (b) kötü huylu tümör örneği (c) iyi huylu tümörün yer gerçekliği (d) kötü huylu tümörün yer gerçekliği)



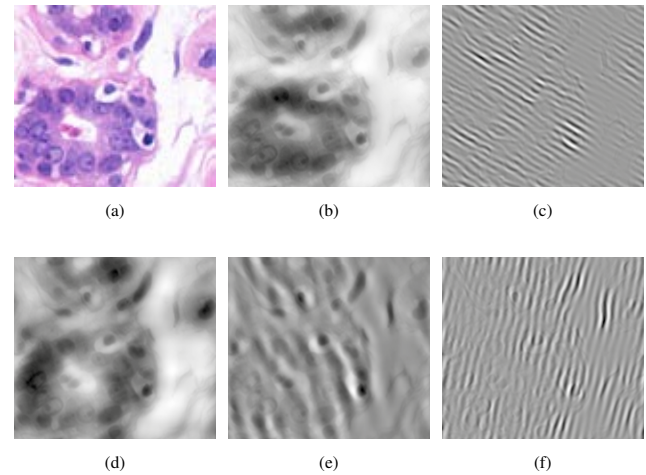
Şekil 2: Uygulama adımlarının genel akış şeması

Sınıflandırma için veriler, Matlab® ile GKA ve DKA yöntemlerinin iki boyutlu verilere uygulanan sürümlerine [20] [16] tabi tutulmuştur. GKA için İKF sayısı 10, yinelemeler arasındaki fark değeri (ϵ) $1e^{-4}$ olarak belirlenmiştir (Şekil 3). DKA için kip sayısı 10, hata değeri $1e^{-7}$, bant genişlik kısıt aralığı [1000 2000] olarak belirlenmiştir (Şekil 4).

Sınıflandırma yöntemi olarak Rassal Ormanlar (RO) [21] ve Destek Vektör Makineleri (DVM) [22], [23] kullanılmıştır. Her iki yöntemde de 20 görüntü eğitim, 38 görüntü de test aşamasında kullanılmıştır. DVM'lerde radyal tabanlı çekirdek fonksiyonlu seçilmiş, ceza parametresi (c) [1-50] aralığında; tekil örneklerin eğitime etkisini gösteren gama (γ) parametresi



Şekil 3: Veri kümesindeki 1 numaralı görüntünün GKA sonuçları ((a) 1 numaralı orijinal görüntü, (b-f) 1 den 5 e kadar olan kipler)



Şekil 4: Veri kümesindeki 1 numaralı görüntünün DKA sonuçları ((a) 1 numaralı orijinal görüntü, (b-f) 1 den 5 e kadar olan kipler)

[0.01-5] aralığında belirli adımlarla artırılarak, en yüksek sınıflandırma başarımları bulunmuştur. Rassal Ormanlarda da yineleme parametresi [100-500] arasında belirli adımlarla artırılmış ve kullanılan özellikler normalizasyon işlemiyle [0-255] aralığına uyarlanarak en iyi sınıflandırma başarımları bulunmuştur. Sınıflandırma başarımları, 10-katlamalı çapraz doğrulama ile hesaplanmış ve *doğruluk*, *kappa değeri* ve *F1-skoru* olmak üzere 3 farklı kistas ile değerlendirilmiştir. Çalışma sonuçları Tablo II'de ve Tablo III'te gösterilmiştir.

RO kullanılarak ÖK-1'de en iyi sınıflandırma başarımları, %87.48 olarak hesaplanmıştır. ÖK-2'de normalizasyon, başarımları etkilememiş ve sınıflandırma başarımları %87.61 olarak elde edilmiştir. RO'daki en yüksek başarımları, ÖK-3 kullanılarak %88.90 ile ulaşılmıştır.

Tablo II: RO kullanılarak elde edilen sınıflandırma sonuçları

Özellik Kümeleri	Rassal Ormanlar		
	Doğruluk (%)	Kappa	F1-Skoru
ÖK-1	87.48	0.49	0.86
ÖK-2	87.61	0.51	0.88
ÖK-3	88.90	0.52	0.89

Tablo III: DVM kullanılarak elde edilen sınıflandırma sonuçları

Özellik Kümeleri	Destek Vektör Makineleri		
	Doğruluk (%)	Kappa	F1-Skoru
ÖK-1	87.54	0.57	0.86
ÖK-2	88.72	0.59	0.88
ÖK-3	89.34	0.60	0.90

DVM'de özellik kümelerinin normalizasyonu başarımı olumlu düzeyde etkilemiştir. Ancak γ değerinin 1.0'dan büyük olması aşırı eğitime sebep olmuş, sınıflandırma başarımını olumsuz etkilemiştir. ÖK-1 ile %87.54'lük, ÖK-2 kümesi ile %88.72'lik ve ÖK-3 kümesi ile de, %89.34'lük sınıflandırma başarımları elde edilmiştir. Doğruluk, Kappa ve F1-skorları incelendiğinde DVM'nin ÖK-3 özellik kümesini kullanarak en iyi sınıflandırma başarımına eriştiği gözlenmiştir.

IV. SONUÇ

Bu çalışmada, histopatolojik kanser görüntülerindeki hücrelerin bölütlenmesi başarımının veri uyumlu özellik çıkarım yöntemleriyle artırılması amaçlanmıştır. Son yıllarda yaygın kullanılan veri uyumlu yöntemlerden olan GKA ve DKA'nın iki boyutlu sürümleri ile özellikler çıkarılmış ve sınıflandırma başarımına katkıları 3 farklı kıstasla (*doğruluk*, *kappa* ve *F1-skoru*) incelenmiştir. Uygulama sonuçları incelendiğinde DKA öznelikleri, GKA özneliklerine ve orijinal piksel değerlerine göre daha yüksek başarımlar sağlamıştır. Ayrıca iki kıstasta da topluluk öğrenme yöntemi olan RO yerine çekirdek öğrenme yöntemi olan DVM daha iyi sınıflandırma başarımına ulaşmıştır. Sonuç olarak, veri-uyumlu bu iki yöntemden elde edilen özneliklerin, orijinal piksel değerleriyle oluşturulan özneliklere göre daha ayrıştırıcı niteliklere sahip olduğu gözlenmiştir.

Bu çalışmanın bir sonraki aşaması olarak veri uyumlu yöntemlerle elde edilen özelliklerin, evrimsel sinir ağları başta olmak üzere derin öğrenme algoritmalarında kullanılması ve sonuçlarının değerlendirilmesi öngörülmektedir.

KAYNAKÇA

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A cancer journal for clinicians*, vol. 68, no. 6, pp. 394–424, 2018.
- [2] H. Mohan, *Textbook of Pathology*. Jaypee Brothers, Medical Publishers Pvt. Limited, 2018.
- [3] C. W. Elston and I. O. Ellis, "Pathological prognostic factors in breast cancer. i. the value of histological grade in breast cancer: experience from a large study with long-term follow-up. CW Elston & IO Ellis. *histopathology* 1991; 19: 403–410," *Histopathology*, vol. 41, no. 3a, pp. 151–151, 2002.
- [4] J. Kong, O. Sertel, H. Shimada, K. L. Boyer, J. H. Saltz, and M. N. Gurcan, "Computer-aided evaluation of neuroblastoma on whole-slide histology images: Classifying grade of neuroblastic differentiation," *Pattern Recognition*, vol. 42, no. 6, pp. 1080–1092, 2009.
- [5] H. Irshad, A. Veillard, L. Roux, and D. Racoceanu, "Methods for nuclei detection, segmentation, and classification in digital histopathology: A review—current status and future potential," *IEEE Reviews in Biomedical Engineering*, vol. 7, pp. 97–114, 2013.
- [6] A. Albayrak and G. Bilgin, "Automatic cell segmentation in histopathological images via two-staged superpixel-based algorithms," *Medical & Biological Engineering & Computing*, vol. 57, no. 3, pp. 653–665, 2019.
- [7] M. N. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Reviews in Biomedical Engineering*, vol. 2, p. 147, 2009.
- [8] N. Hatipoğlu and G. Bilgin, "Segmentation of histopathological images with convolutional neural networks using fourier features," in *IEEE 23rd Signal Processing and Communications Applications Conference, SIU'15*. IEEE, 2015, pp. 455–458.
- [9] N. Hatipoğlu and G. Bilgin, "Cell segmentation in histopathological images with deep learning algorithms by utilizing spatial relationships," *Medical & Biological Engineering & Computing*, vol. 55, no. 10, pp. 1829–1848, 2017.
- [10] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proc. of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971. The Royal Society, 1998, pp. 903–995.
- [11] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 01, pp. 1–41, 2009.
- [12] Y. Yang, J. Deng, and C. Wu, "Analysis of mode mixing phenomenon in the empirical mode decomposition method," in *IEEE 2nd Int. Symp. on Information Science and Engineering*, 2009, pp. 553–556.
- [13] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.
- [14] Y. Wang, R. Markert, J. Xiang, and W. Zheng, "Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system," *Mechanical Systems and Signal Processing*, vol. 60, pp. 243–251, 2015.
- [15] W. Liu, S. Cao, and Y. Chen, "Applications of variational mode decomposition in seismic time-frequency analysis," *Geophysics*, vol. 81, no. 5, pp. V365–V378, 2016.
- [16] K. Dragomiretskiy and D. Zosso, "Two-dimensional variational mode decomposition," in *Int. Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, 2015, pp. 197–208.
- [17] S. Lahmiri, "Image characterization by fractal descriptors in variational mode decomposition domain: application to brain magnetic resonance," *Physica A: Statistical Mechanics and its Applications*, vol. 456, pp. 235–243, 2016.
- [18] S. Chatteraj and K. Vishwakarma, "Classification of histopathological breast cancer images using iterative VMD aided Zernike moments & textural signatures," *arXiv preprint arXiv:1801.04880*, 2018.
- [19] E. D. Gelasca, B. Obara, D. Fedorov, K. Kvilekval, and B. Manjunath, "A biosegmentation benchmark for evaluation of bioimage analysis methods," *BMC Bioinformatics*, vol. 10, no. 1, p. 368, 2009.
- [20] A. Linderhed, "Image empirical mode decomposition: A new tool for image processing," *Advances in Adaptive Data Analysis*, vol. 1, no. 02, pp. 265–294, 2009.
- [21] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [22] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [23] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.