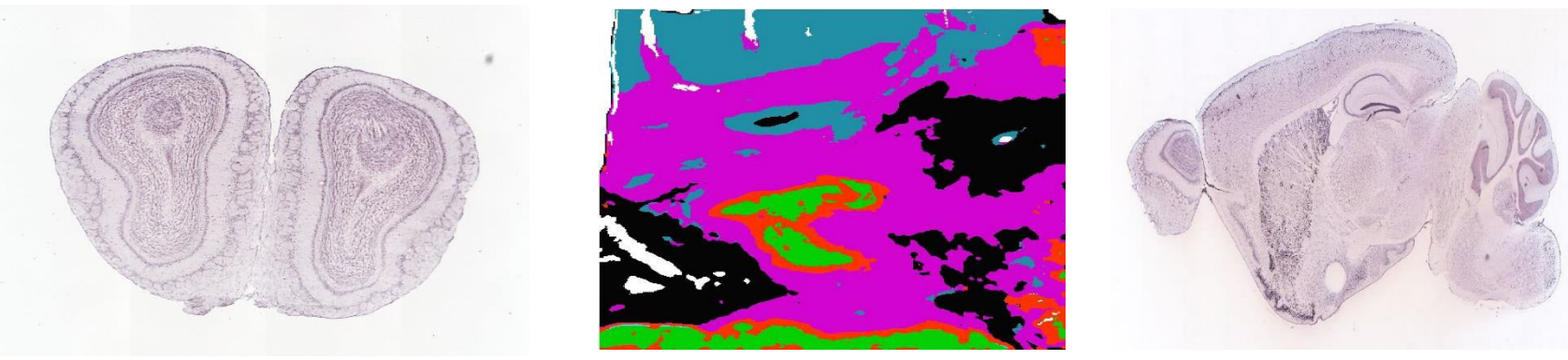


Özet

Çalışma biyoenformatik alanında gen yapılarını tanımlamada bilim insanlarına yardımcı olmayı hedeflemektedir. Bu noktadan hareketle derin öğrenme teknikleri olan otomatik kodlayıcılar (auto encoders-AE) ve evrimsel yapay sinir ağlar (CNN) kullanılarak sinir hücrelerinin resimlerinden elde edilen özelliklerin gen ontolojisindeki (GO) hangi işlevsellik kategorisine ait olduğu belirlenerek sinir resimleri üzerinden sınıflandırma yapılabilmesi hedeflenmektedir.



Auto Encoders

CNN

İşlevselliğine dair özelliklerinin temsili

Benzerliklerin karşılaştırılarak yeni gen tanımlarının keşfi (SVM)



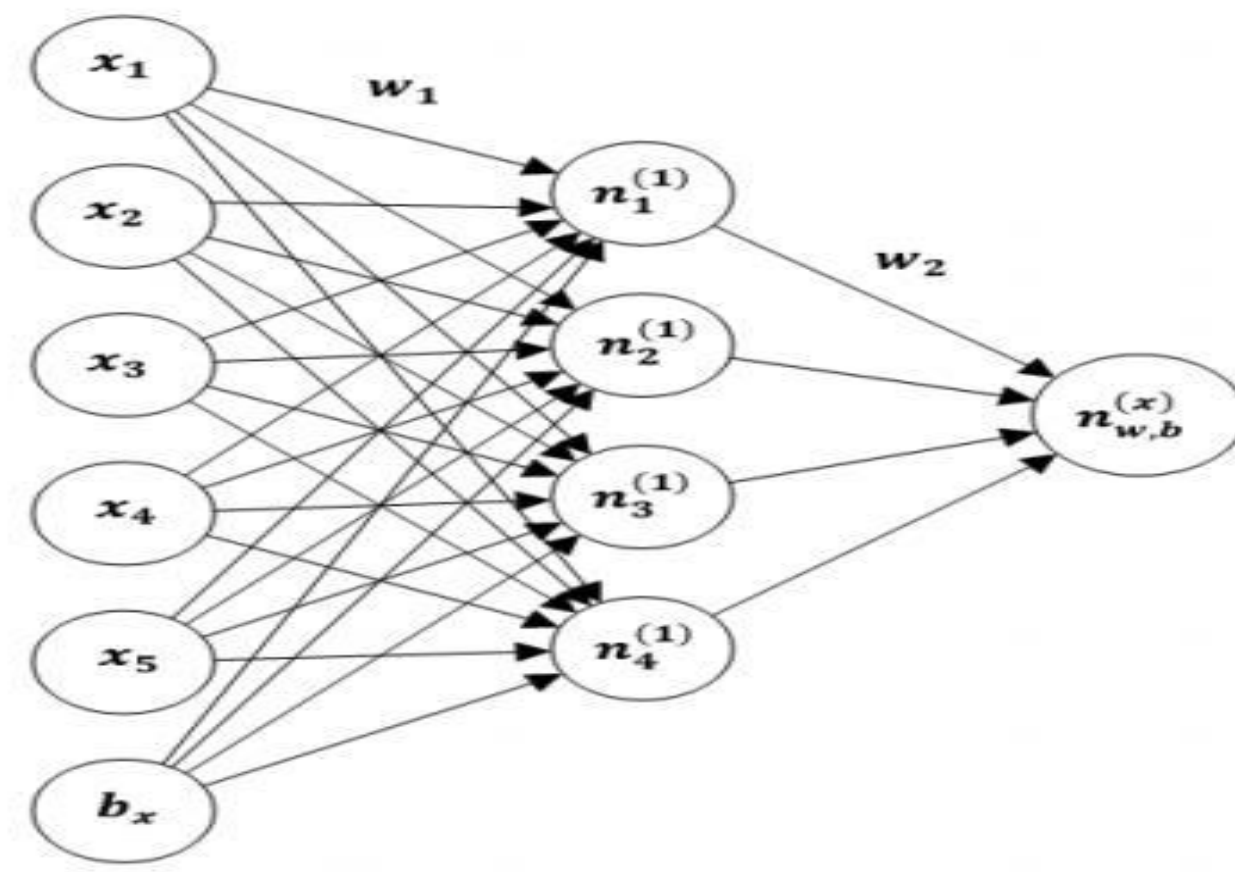
Şekil 1. Akış Şeması

Giriş

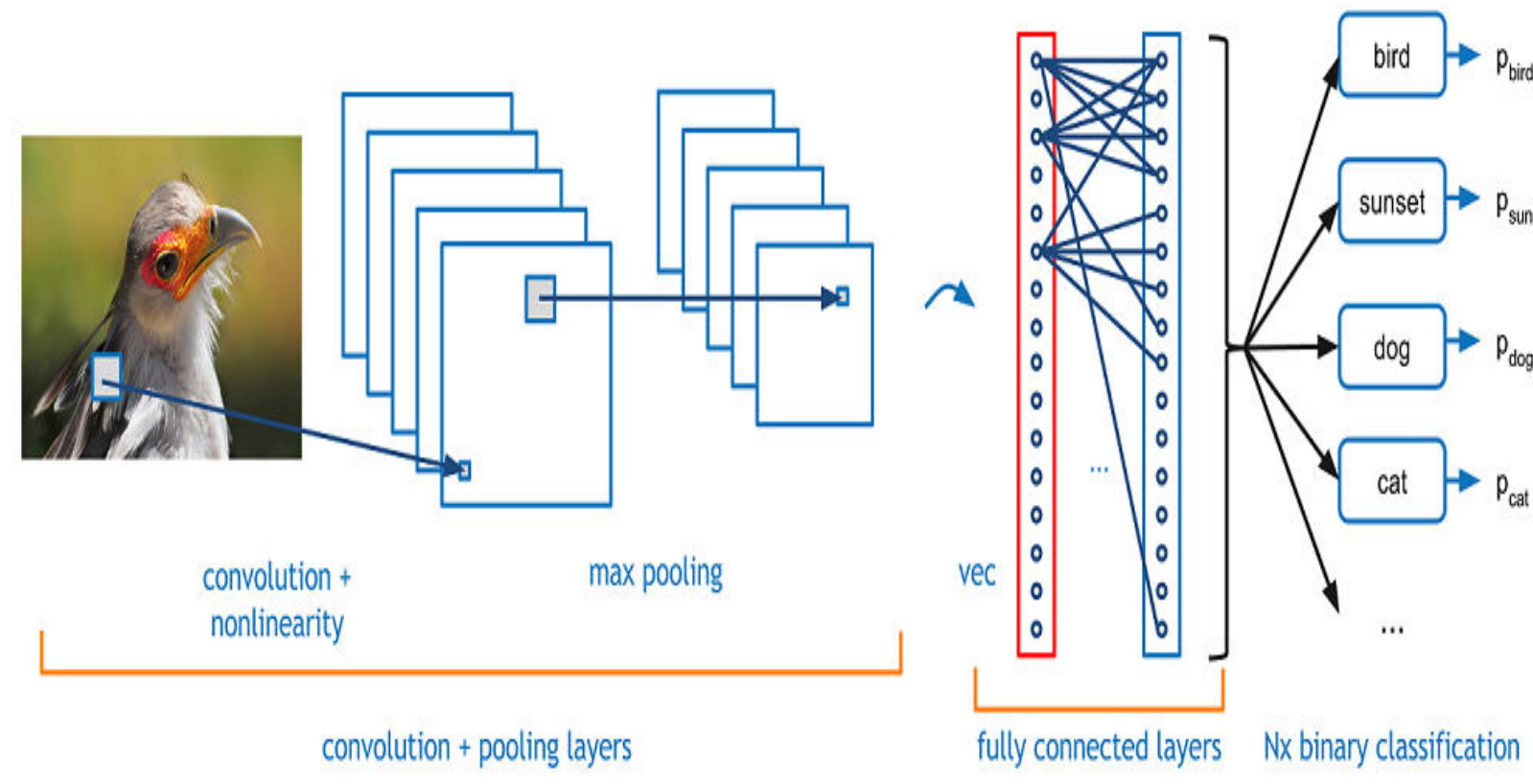
Nöronların çok fazla miktardaki değişken çözünürlükteki resimleri üzerinde analizler yapılarak gen yapılarına dair bilgiler elde edilebilmektedir. Örneğin *in situ hybridization* (ISH) olarak adlandırılan memeli beyin resimlerinden [1] oluşan veri seti, nükleik asitlerin belirli hücrelerdeki yerini saptamak adına genlerin anlatımında etkin bir yaklaşım sergilemeye imkan tanımaktadır [2]. Bu noktada resimlerden ilgili işlevselliği yakalayarak sinir bilimcilerin yorum yapabileceği anlamlı bir deseni ortaya koymak gerekmektedir. Bunun yararlarından biri elde edilen desen ile sınıflandırma yapılarak Gen Ontolojisinde (GO) hangi kategoriye girdiğini anlamaktır [3]. Yeni gen işlevselliği tanımlayıcı bilgilerini resimlerden elde edilen desen ve GO terimleri arasındaki bağlantıyı çözerek sunmak bu alanda çok değerli katkılar sağlayacaktır [4].

Yöntem

Sinir resimlerinden özellik çıkarmadaki ana amaçlardan birisi uzmanlar tarafından dahi zor görülebilecek noktaları ortaya çıkarmaktır. Son zamanlarda GO 'ya yeni eklemelerde bulunabilmek için makine öğrenimi kullanılmaktadır [2,5,6]. Bu çalışmada derin öğrenme metotları: AE ve CNN kullanılmış ve bunun yanında AE metodu Taguchi tekniği ile desteklenmiş [7] ayrıca benzerlikler tespit edilirken karar destek makinesi (SVM) ile birlikte boyut azaltma (PCA) [8] kullanılmıştır. Çalışmanın başarımı standart ölçütlerle değerlendirilmiş önceki yöntemlerin başarımları ile kıyaslanmıştır.

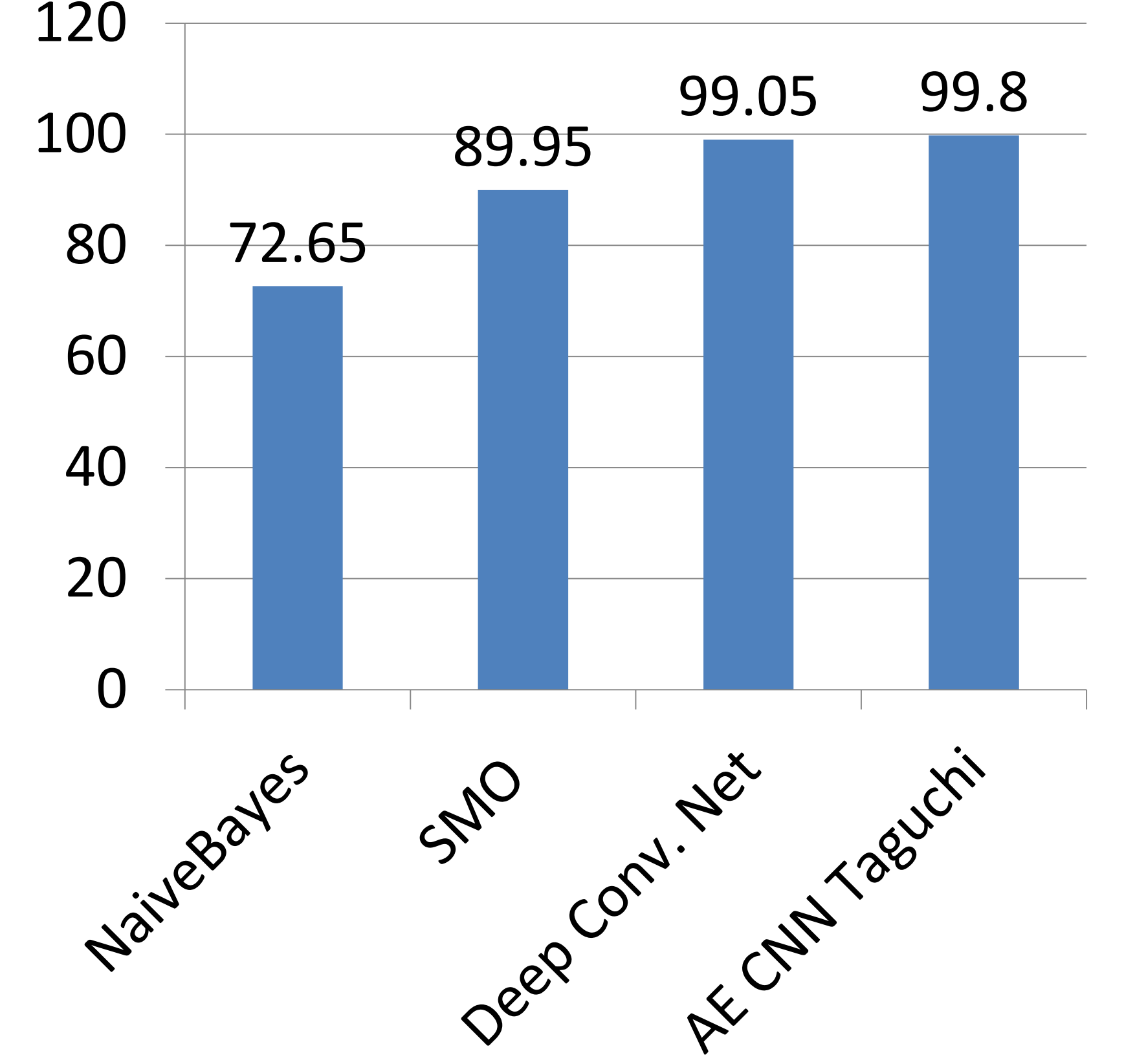


Şekil 2. Basit auto encoder gösterimi.



Şekil 3. CNN Örnek gösterimi

Sonuç



Grafik 1 . Performans sonuçlarının karşılaştırılması.

Tartışma

Bu çalışmada önemli noktalardan birisi yerel bir beyin bölgesinin histogramı üzerinden tanımlayıcılar ile gen ontolojisinde karşılık gelen işlevselliği yakalamanın [9] aksine genel bir yöntemin derin öğrenme ile geliştirilerek herhangi bir beyin bölümü resimleri üzerinden de uygulanabilmesidir.

AE' nin faydası özellik çıkarımının otomatikleşmesini sağlama ve buradan elde edilen özellikler ile CNN ' e girdi oluşturup sınıflandırmayı sağlayacak çıktılarını elde edilmesidir. Aynı zamanda taguchi metodu ile özellikle zaman maliyetinden kazanım sağlanmıştır.

SVM ise PCA ile birlikte kullanılarak yine doğruluktan fazla bir kayıp olmaksızın performanslı çalışmayı sağlamıştır.

Sonuç

ISH verileri sinir görüntüleri üzerinden ilgili beyin bölümünde hangi tür genler yer aldığını anlamak açısından biyoloji ve tıp alanında büyük önem arz etmektedir. Beyin kesitlerinden farklı açı, konum ve çözünürlükten elde edilen resimlerin analizi zor ve zaman alan süreçlerdir. Neyse ki derin öğrenmenin bu analizleri kolaylaştırması ile bu bilim dalına da çok önemli katkılarda bulunulmuştur.

Bu çalışmada da hem derin öğrenme, hem anlamsal ağlar ve de makine öğrenimi kullanılarak beyin sinir resimlerinden elde edilen bulgular gen ontolojisi ile karşılaştırılmış ve sonuçta işlevsel benzerliklerine göre doğru şekilde sınıflandırılması sağlanmıştır. Böylece genlerin tanımlanması, ve anlamlandırılması noktasında bilime önemli katkılarda bulunulmuştur.

Parametre Sayısı

	2	3	4	5	6	7	8	9
2	L4	L4	L8	L8	L12	L12	L12	L16
3	L9	L9	L9	L18	L18	L18	L27	L27
4	L16	L16	L16	L16	L32	L32	L32	L32
5	L25	L25	L25	L25	L25	L50	L50	L50

Tablo 1. Taguchi Metodu Seviye ve parametre sayısı ilişkisi

İletişim

Ömer SEVİNÇ
 Ondokuz Mayıs Üniversitesi
 Sevinc.omer@gmail.com
<https://personel.omu.edu.tr/tr/osevinc>
 Telefon: 03626470246

Referanslar

- Henry, A. M., & Hohmann, J. G. (2012). High-resolution gene expression atlases for adult and developing mouse brain and spinal cord. *Mammalian genome*, 23(9-10), 539-549.
- Cohen, I., David, E. O., Netanyahu, N. S., Liscovitch, N., & Chechik, G. (2017, September). DeepBrain: Functional Representation of Neural In-Situ Hybridization Images for Gene Ontology Classification Using Deep Convolutional Autoencoders. In *International Conference on Artificial Neural Networks* (pp. 287-296). Springer, Cham.
- Gene Ontology Consortium. (2007). The gene ontology project in 2008. *Nucleic acids research*, 36(suppl_1), D440-D444.
- Du Plessis, L., Škunca, N., & Dessimoz, C. (2011). The what, where, how and why of gene ontology—a primer for bioinformaticians. *Briefings in bioinformatics*, 12(6), 723-735.
- Pinoli, P., Chicco, D., & Masseroli, M. (2015). Computational algorithms to predict Gene Ontology annotations. *BMC bioinformatics*, 16(6), S4.
- Kordmahalleh, M. M., Homafar, A., & Dukka, B. K. C. (2013, November). Hierarchical multi-label gene function prediction using adaptive mutation in crowding niching. In *Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on* (pp. 1-6). IEEE.
- Karim, A. M., Güzel, M. S., Tolun, M. R., Kaya, H., & Çelebi, F. V. (2018). A New Generalized Deep Learning Framework Combining Sparse Autoencoder and Taguchi Method for Novel Data Classification and Processing. *Mathematical Problems in Engineering*, 2018.
- Luo, Y., Wu, C. M., & Zhang, Y. (2013). Facial expression recognition based on fusion feature of PCA and LBP with SVM. *Optik-International Journal for Light and Electron Optics*, 124(17), 2767-2770.
- Liscovitch, N., Shalit, U., & Chechik, G. (2013). FuncISH: learning a functional representation of neural ISH images. *Bioinformatics*, 29(13), i36-i43.