

Etkili Çeltik Hastalığı Tespiti için Vision Transformer Modellerinin Kullanımı

Cüneyt ÖZDEMİR^{1*}

¹Bilgisayar Mühendisliği / Mühendislik Fakültesi, Siirt Üniversitesi, Türkiye

*(cozdemir@siirt.edu.tr)

Özet – Bu araştırma, Vision Transformer modellerini kullanarak çeltik hastalıklarının otomatik olarak tespit edilmesine odaklanmaktadır. Çeltik, temel bir gıda kaynağı olarak büyük bir öneme sahiptir ve çeltik üretimine olan talep giderek artmaktadır. Bu nedenle, etkili hastalık tespit yöntemlerine ihtiyaç duyulmaktadır. Çeltik ürünlerinde hastalık tespiti için kullanılan manuel teknikler zaman alıcı olabilir ve hatalara açık olabilir. Ancak, otomatik tespit sistemleri, sensörler ve makine öğrenimi algoritmalarını kullanarak daha doğru ve maliyet açısından etkili bir yaklaşım sunar. Bu çalışmada, Paddy Doctor veri kümesi kullanılarak 13 farklı çeltik hastalığına ait 4.160 görüntü üzerinde çalışma yapılmıştır. Doğru hastalık sınıflandırması yapmak için dört farklı vision transformer modeli kullanılmıştır. Deneysel sonuçları, ViT-B16 modelinin yaklaşık %92,88 doğruluk oranıyla en yüksek başarıyı gösterdiğini ortaya koymuştur. Bu bulgular, Vision Transformer modellerinin çeltik hastalıklarını etkili bir şekilde tespit etmek için kullanılabilirliğini göstermektedir. Bu durum, çeltik üretiminde mahsul verimini artırarak pestisit kullanımını azaltmaya ve çevre korumasını iyileştirmeye katkı sağlayabilir. Ayrıca, Vision Transformer modellerinin otomatik hastalık tespitindeki potansiyelini ortaya koymakta ve tarım verimliliği ile çevresel sürdürülebilirlik üzerinde olumlu etkileri olabileceğini göstermektedir.

Anahtar Kelimeler – Çeltik Hastalıkları, Vision Transformer, ViT, Paddy Doctor, Makine öğrenmesi

I. GİRİŞ

Çeltik, küresel nüfusun çoğunluğu için temel bir gıda kaynağı olarak büyük önem taşımaktadır. Dünya nüfusundaki artışın mevcut gidişatı göz önüne alındığında, çeltiğe yönelik artan talebi karşılamak için çeltik üretiminin 2030 yılına kadar %50 oranında artırılması gerektiği öngörülmektedir. Özellikle Çin ve Hindistan, Endonezya, Bangladeş ve Vietnam ile birlikte çeltik üretiminde öne çıkan ülkelerdir. Küresel ortalama çeltik verimi 410 kg/da (dekar) iken, Türkiye 780 kg ile nispeten daha yüksek bir ortalama çeltik verimi sergilemektedir. Türkiye'de çeltik ekimi, Edirne, Balıkesir, Çorum, Samsun, Sinop ve Kastamonu gibi bölgeler başta olmak üzere 31 ilde yapılmaktadır. Çeltik, yaklaşık %5-10 oranında protein içeriğine sahiptir ve beslenme için gerekli olan zengin amino asit bileşimi nedeniyle

buğdaydan sonra oldukça rağbet gören bir üründür. Çeltik yetiştiriciliğinde toprak seçimi için kesin bir kriter bulunmamasıyla birlikte, en uygun koşullar besin maddesi bakımından bol, derin, tınlı ve su geçirgenliği düşük topraklarda bulunur. Buna karşılık, kumlu topraklar su kaybı ve besin eksikliği eğilimleri nedeniyle çeltik ekimi için uygun değildir [1].

Çeltik, bir tahıl ürünü olarak yüksek verimliliği nedeniyle hem yurt içinde hem de yurt dışında tercih edilen bir temel tarım ürünüdür. Önemli ölçüde getiri sağlama kapasitesi nedeniyle üreticiler için önemli bir seçeneği temsil etmektedir. Küresel nüfus artışı, artan refah ve gelişen alışkanlıklar ile birleştiğinde, Türkiye ve diğer bazı ülkelerde çeltik tüketiminde bir artış eğilimine yol açmıştır [2].

Çeltik yetiştiriciliği, karmaşık bir süreç olup çeşitli hastalık ve zararlılara yatkındır. Bu

hastalıkların ve zararlıların erken teşhisi, çiftçiler ve tarım uzmanları için son derece önemli bir görev haline gelmiştir. Geleneksel olarak, çiftçiler çeltik hastalıklarının tespit etmek için deneyimlerine ve görsel incelemelerine dayanan manuel tekniklere güvenirliler. Ancak bu teknikler son derece verimsiz, zaman alıcı ve hataya açıktır. Deneyimli çiftçiler ve tarım uzmanları bile benzer hastalık belirtilerinin çokluğu nedeniyle mahsul hastalıklarını doğru bir şekilde tespit etmekte zorlanabilmektedir [3-5].

Çeltik hastalıklarının otomatik olarak tespit edilmesi, hastalıkların tanımlanması ve yönetilmesinde daha etkili, doğru ve maliyet tasarruflu bir yaklaşım sunarak bu zorlukların üstesinden gelmeyi sağlayabilir. Otomatik tespit sistemleri, çeltik bitkileriyle ilgili verileri toplamak için telefon kameraları, standart kameralar, spektrometreler ve termal sensörler gibi çeşitli sensörleri kullanabilir. Bu veriler daha sonra, hastalıkların erken aşamada tespit edilmesini sağlamak için makine öğrenimi algoritmalarıyla analiz edilebilir. Çeltik hastalıklarının otomatik tespiti, mahsul verimini artırma, pestisit kullanımını azaltma ve çevrenin korunması gibi potansiyel faydalar sunmaktadır. Son dönemde, makine öğrenme modellerinin çeşitli hastalıkların tespitinde kullanımı artmıştır ve bu modeller, farklı problemlerde yüksek doğruluk sağlama yetenekleriyle dikkat çekmektedir [6-10].

Bu çalışmada, makine öğrenmesi teknolojilerinden Vision Transformer (ViT) modelleri kullanılarak 13 farklı çeltik hastalığının tespiti yaklaşık %92.88 oranında bir doğrulukla gerçekleştirilmiştir.

II. MATERYAL VE METOT

A. Materyal

Bu çalışmada, 13 farklı çeltik hastalığına ilişkin verileri içeren "paddy doctor" veri kümesi [11] kullanılmıştır. Hastalık grubu başına 320 görüntü olmak üzere toplam 4.160 görüntü kullanılmıştır. Sınıflar ve bunlara karşılık gelen miktarlar Tablo 1'de sunulmuştur. Veri toplama süreci, çeltik bitkilerinin 40 ila 80 günlük olduğu Şubat ve Nisan 2021 tarihleri arasında gerçekleştirilmiştir. Her görüntü, hastalık belirtilerinin varlığına dayalı olarak bir tarım uzmanı ile işbirliği içinde manuel etiketlemeye tabi tutuldu ve bir hastalık sınıfı etiketi atandı. Etiketlenen çeltik hastalıkları arasında Bakteriye Yaprak Yanıklığı (BLB), Bakteriye

Yaprak Çizgisi (BLS), Bakteriye Salkım Yanıklığı (BPB), Kök Kurdu (BSB), Patlama, Kahverengi Leke, Pas, Hispa, Yaprak Küfü, Tungro, Beyaz Kök Kurdu, Sarı Kök Kurdu ve Normal Yaprak yer almaktadır [11].

Tablo 1. Paddy doctor veri seti

Hastalık	Eğitim Sayısı	Test Sayısı
Tungro	320	80
Sarı Kök Kurdu (Yellow Stem Borer)	320	80
Hispa	320	80
Tüylü Küf (Downy Mildew)	320	80
Siyah Gövde Kurdu (Black Stem Borer)	320	80
Bakteriye Yaprak Çizgisi (Bacterial Leaf Streak)	320	80
Bakteriye Yaprak Yanıklığı (Bacterial Leaf Blight)	320	80
Kahverengi Leke (Brown Spot)	320	80
Patlamış (Blast)	320	80
Yaprak Silindiri (Leaf Roller)	320	80
Normal	320	80
Bakteriye Salkım Yanıklığı (Bacterial Panicle Blight)	320	80
Beyaz Gövde Kurdu (White Stem Borer)	320	80
Toplam	4160	1040

B. Metot

Vision Transformer (ViT), görüntülerin temsillerini öğrenmek için self-attention mekanizmasını kullanan bir sinir ağı mimarisidir. Alexey Dosovitskiy ve diğerlerinin [12] önerdiği bir çalışmada ilk kez tanıtılmıştır. ViT, başlangıçta doğal dil işleme (NLP) görevleri için geliştirilen transformatör mimarisine dayanmaktadır. Görüntü tanıma için kullanılan bir yapay sinir ağı türüdür.

ViT, bir görüntüyü bir dizi görüntü yaması olarak temsil eder ve bu yamalar arasındaki ilişkileri öğrenmek için self-attention mekanizmasını kullanır. Self-attention, herhangi bir yama çiftinin görüntüdeki herhangi bir konumda birbirleriyle etkileşime girebilmesini sağlayan bir mekanizmadır. Bu, geleneksel görüntü işleme yaklaşımlarından farklıdır, çünkü ViT'nin dikkate aldığı özelliklerin konumlarına bağlı olmadığı için yerel bölgelere sınırlı kalmaz.

ViT modelleri, görüntü tanıma görevlerinde son teknoloji ürünü sonuçlar elde edebilme potansiyeline sahiptir. Bu modeller, büyük veri kümeleri üzerinde eğitildiğinde etkileyici performans göstermektedir. ViT, görüntü işlemede geleneksel olarak kullanılan Convolutional Neural Networks (CNN'ler) yerine, dikkat mekanizmasının gücünü kullanarak daha esnek bir yaklaşım sunmaktadır.

ViT'nin farklı çeşitleri bulunmaktadır. Orijinal ViT modeli, görüntüleri 16x16 görüntü yamaları

olarak temsil ederken, diğer türler farklı yama boyutları ve gizli katman sayılarına sahiptir. ViT-B32, en küçük modeldir ve 110 milyon parametreye sahiptir. ViT-L16 ise en büyük modeldir ve 340 milyar parametreye sahiptir. Model boyutu arttıkça, daha fazla parametreye sahip olur ve daha karmaşık temsiller öğrenebilir. Ancak, daha büyük modellerin eğitimi için daha fazla veri ve hesaplama kaynakları gerekmektedir. Tablo 2'de ViT model türlerini özetleyen bir tablo verilmiştir.

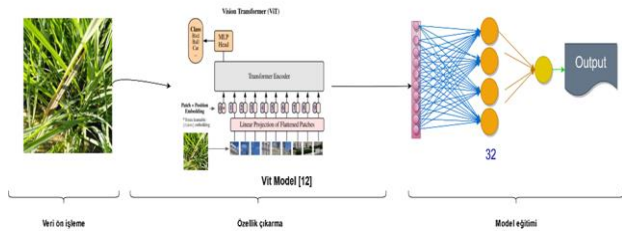
Tablo 2. ViT Model türleri

Model Name	Parametre Sayısı	Yama boyutu	Gizli Katman sayısı
ViT-B32	110 milyon	16x16	128
ViT-B16	340 milyon	32x32	128
ViT-L32	1.3 milyar	64x64	384
ViT-L16	340 milyar	128x128	384

Bu çalışma, Paddy Doctor veri setinin özgün özelliklerine uyacak şekilde ViT modellerini adapte etmeyi amaçlamıştır. Adaptasyonun ardından, bu özelleştirilmiş modellerin performansı titizlikle incelenmiş ve veri kümesi üzerinde analiz edilmiştir.

Model mimarisi, Şekil 1'de gösterildiği şekilde bu çalışma için kullanılmıştır. İlk olarak, veri kümesi ön işlemden geçirilir. Ardından, ViT modelleri kullanılarak özellikler çıkarılır. Son olarak, çıkarılan özellikler düzleştirilir (flatten) ve 32 nöronlu ve 13 çıkışa sahip bir yapay sinir ağına giriş olarak verilir.

Bu yöntem, Paddy Doctor veri setinin özel gereksinimlerine uygun olarak ViT modellerini kullanırken, özellik çıkarımı ve sınıflandırma için bir yapay sinir ağı kullanır. Bu yaklaşımın performansı, dikkatli bir şekilde incelenmiş ve veri kümesi üzerinde yapılan analizlerle değerlendirilmiştir.



Şekil 1. Model Mimarisi

Veri çeşitliliğini artırmak için 10 derece döndürme, %20 yatay ve dikey kaydırma, %20

yakınlaştırma veya uzaklaştırma, %5 yatay kaydırma, %5 dikey kaydırma, görüntüleri yatay olarak simetriye göre döndürme, Piksel değerlerini 0 ile 1 arasında ölçeklendirme işlemi yapmıştır. Ayrıca özellikle ViT-L modellerinin çok büyük parametre barındırmasından kaynaklı olarak görüntü boyutları 192x192 olacak şekilde yeniden boyutlandırılmış ve batch size parametresi 16 olarak ayarlanmıştır.

III. BULGULAR

Veri seti, ön işlemden geçirildikten sonra eğitim (train), doğrulama (validation) ve test olarak ayrılmıştır. Görüntülerin %75'i eğitim, %25'i ise test veri kümesini oluşturacak şekilde ayrılmıştır. Eğitim veri kümesi, 13 sınıfa ayrılan toplam 4160 görüntüden her sınıfta 320 görüntü içermektedir. Bu görüntülerin %15'i, yani 624 tanesi doğrulama veri kümesi için ayrılmıştır. Geriye kalan 3536 görüntü ise eğitim için kullanılmaktadır. Test veri kümesi ise her sınıfta 80 görüntü olacak şekilde 13 sınıfa ayrılan toplam 1040 görüntüden oluşmaktadır.

Veri setinin ön işlemleri tamamlandıktan sonra çeltik hastalıklarının sınıflandırılması için ViT modelleri kullanılmıştır.

Bu çalışmada, çeltik hastalıklarını sınıflandırmak için kullanılan tüm modellerde optimizasyon algoritması olarak Adam optimizasyon algoritması kullanılmış ve öğrenme oranı 0,001 olarak ayarlanmıştır. Her model 100 epokta eğitilmiştir. ModelCheckpoint ve EarlyStopping yöntemleri de kullanılmış ve modelin en iyi performans gösterdiği aşamada eğitim durdurulmuştur.

Veri seti üzerinde gerçekleştirilen deneylerde ViT modellerinin performansı incelenmiş ve sonuçları Tablo 2'de sunulmuştur. Tablo, farklı ViT modellerinin çeltik hastalıklarını sınıflandırmada elde ettikleri sonuçları içermektedir. Bu çalışma, modellerin doğruluk, hassasiyet, özgünlük veya F1 skoru ölçütlerine göre performanslarını analiz etmiştir.

Tablo 3. ViT modelleri performans sonuçları

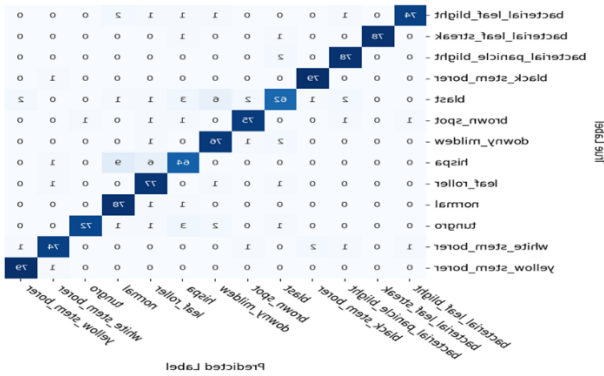
	Accuracy	F1 score	Recall	Precision
ViT-B32	%90,56	%89,49	%89,56	%89,84
ViT-B16	%92,88	%92,83	%92,88	%93,05
ViT-L32	%89,08	%89,05	%89,07	%89,24
ViT-L16	%83,98	%83,97	%83,98	%84,81

Tablo 3, ViT-B16 modelinin %92,88'lik doğruluk oranıyla en iyi sonuçları elde ettiğini göstermektedir.

- ViT-L16 modeli en düşük doğruluğa, F1 puanına, hassasiyete ve kesinliğe sahiptir. Bu, modelin görüntüleri sınıflandırmada diğer modeller kadar iyi olmadığını göstermektedir.
- ViT-B16 modeli en yüksek doğruluk, F1 puanı, kesinlik ve hassasiyete sahiptir. Bu, modelin görüntüleri sınıflandırmada en iyi olduğunu göstermektedir.
- ViT-B32 ve ViT-L16 modelleri en iyi doğruluk, F1 puanı, kesinlik ve hassasiyete sahiptir. Bu, modellerin görüntüleri sınıflandırmada iyi olduğunu göstermektedir.

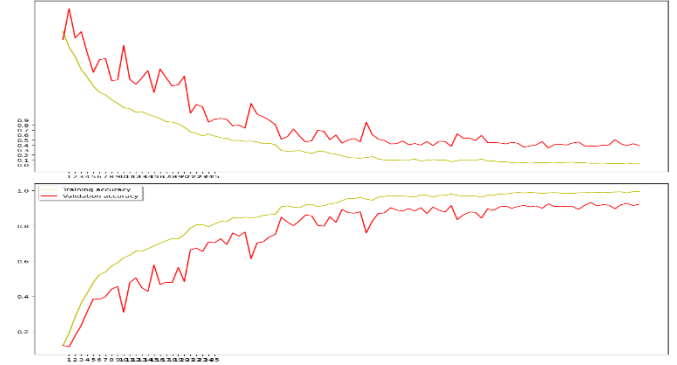
Genel olarak, ViT-B16 modeli görüntüleri sınıflandırmada en iyisidir, onu ViT-B32 ve ViT-L32 modelleri takip eder. ViT-L16 modeli, görüntüleri sınıflandırmada en kötü olanıdır.

ViT-B16 modelinin test veri setinde doğru ve yanlış sınıflandırdığı görüntülerin sayısını gösteren karmaşıklık matrisi Şekil 2'de sunulmuştur. Bu matris, modelin her bir sınıf için ne kadar doğru tahmin yaptığını ve ne kadar yanlış tahmin yaptığını gösteren bir karmaşıklık matrisidir.



Şekil 2. ViT-B16 modeline ait karmaşıklık matrisi

Şekil 3'te en iyi performansı gösteren ViT-B16 modelinin kayıp(loss) ve doğruluk (accuracy) grafiği gösterilmiştir.



Şekil 3. ViT-B16 modelinin kayıp ve doğruluk grafiği

Tablo 4, çeltik hastalıklarını sınıflandırmak için ViT-B16 modelinin kullanıldığı ve 13 farklı sınıfa ayrılan görüntüler üzerinde elde edilen sınıflandırma sonuçlarını göstermektedir. Bu tablo, her sınıfa ait görüntülerin, farklı metrikler kullanılarak başarılarını yüzdelik olarak göstermektedir. Bu metrikler, doğruluk, hassasiyet, özgünlük, F1 skoru ölçütlerini ve her sınıfa ait test görüntü sayısını içermektedir.

Tablo 4. ViT-B16 Sınıflandırma raporu

Class	Precision	Recall	F1-Score	Support
bacterial_leaf_blight	0.973	0.925	0.949	80
bacterial_leaf_streak	1.000	0.975	0.987	80
bacterial_panicle_blight	0.939	0.975	0.957	80
black_stem_borer	0.963	0.988	0.975	80
blast	0.899	0.775	0.832	80
brown_spot	0.949	0.938	0.943	80
downy_mildew	0.884	0.950	0.916	80
hispa	0.865	0.800	0.831	80
leaf_roller	0.865	0.963	0.911	80
normal	0.857	0.975	0.912	80
tungro	0.986	0.900	0.941	80
white_stem_borer	0.948	0.925	0.937	80
yellow_stem_borer	0.963	0.988	0.975	80
Overall				
Accuracy	0.929	1.000	0.964	1040
Macro Avg	0.930	0.929	0.928	1040
Weighted Avg	0.930	0.929	0.928	1040

Tablo 4'te ViT-B16 modelinin, %100 kesinlik, %97,5 hassasiyet ve %98,73 F1 skoru ile bakteriyel_leaf_streak sınıfında en iyi performansı göstermiştir. ViT-B16 modeli, %89,86 kesinlik, %77,5 hassasiyet ve %83,22 F1 puanı ile blast sınıfında en kötü performansı göstermiştir. Genel olarak, model, %92.89'luk bir genel doğrulukla, sınıflandırma görevinde iyi performansı

göstermiştir. Model özellikle bakteriyel_leaf_streak sınıfında iyi performans gösterdi, ancak blast sınıfında geliştirilebilir bir sonuç elde etmiştir.

IV. SONUÇLAR VE TARTIŞMA

Bu çalışmada, ViT modellerini kullanarak çeltik hastalıklarının otomatik tespiti için bir mimari önerildi. ViT modelleri, 13 farklı hastalığa sahip çeltik bitkilerinin 4.160 görüntüsünden oluşan bir veri kümesi üzerinde eğitildi. ViT-B16 modeli %92,88'lik doğruluk oranıyla en iyi performansı elde etti. ViT-L16 modeli %83,98'lik doğruluk oranıyla en kötü performansı elde etmiştir. Bu sonuçlar, çeltik hastalıklarının otomatik tespiti için ViT modellerinin potansiyelini göstermektedir. Sonuçlar, manuel tekniklere daha verimli ve güvenilir bir alternatif sağlayarak, çeltik hastalıklarını doğru bir şekilde tanımlama ve sınıflandırmada ViT modellerinin potansiyelini vurgulamaktadır. Bulgular, hastalık yönetimi ve bitki koruma için umut verici bir yaklaşım sunarak tarım teknolojisi alanına katkıda bulunmaktadır. Hastalıkların erken tespiti, çiftçiler ve tarım uzmanları tarafından zamanında müdahale edilmesini sağlayarak önemli ürün kaybı riskini azaltır. Ayrıca, otomatik algılama sistemlerinin uygulanması, mahsul veriminin artmasına, pestisit kullanımının azalmasına ve çevresel sürdürülebilirliğin iyileştirilmesine yol açabilir.

Bu çalışmanın bazı kısıtlamaları:

* Bu çalışmada kullanılan veri seti nispeten küçüktür. ViT modellerinin performansını daha da iyileştirmek için daha büyük bir veri setine ihtiyaç duyulmaktadır.

* ViT modelleri belirli bir çeltik hastalıkları veri kümesi üzerinde eğitilmiştir. Modeller diğer çeltik hastalıkları veya bitki türleri veri kümelerinde aynı performansı göstermeyebilir.

Bu kısıtlamalara rağmen, bu çalışmanın sonuçları ViT modellerinin çeltik hastalıklarının otomatik tespiti için potansiyelini göstermektedir. Daha fazla araştırma, veri setini genişletmeye ve belirli hastalıkları tespit etme performanslarını artırmak için ViT modellerini iyileştirmeye odaklanabilir. Genel olarak sonuçlara bakıldığında, ViT modelleri gibi makine öğrenimi teknolojilerinin kullanımı, hassas tarımı ilerletmek ve küresel gıda güvenliğini sağlamak için büyük katkılar sağlayabilecek potansiyelde olduğu görülmektedir.

KAYNAKLAR

- [1] Anonim 2007. Çeltik Hastalık ve Zararlıları ile Mücadele. T.C. Tarım ve Orman Bakanlığı Yayınları. URL: https://www.tarimorman.gov.tr/GKGM/Belgeler/Uretici_Bilgi_Kosesi/Dokumanlar/celtik.pdf (erişim tarihi: 16.05.2023)
- [2] Taşlıgil, N., Şahin, G. (2011). Türkiye’de çeltik (*Oryza sativa* L.) yetiştiriciliği ve coğrafi dağılımı. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 4(6), 182-204.
- [3] Amritha Haridasan, Jeena Thomas, and Ebin Deni Raj. 2023. Deep learning system for paddy plant disease detection and classification. *Environmental Monitoring and Assessment* 195, 1 (2023),1–28.
- [4] B. Leelavathy and Ram Mohan Rao Kovvur. 2020. Prediction of Biotic Stress in Paddy Crop Using Deep Convolutional Neural Networks. In *Proceedings of International Conference on Computational Intelligence and Data Engineering*. Springer Singapore, 337–346. https://doi.org/10.1007/978-981-15-8767-2_29
- [5] Vimal K Shrivastava, Monoj K Pradhan, Sonajharia Minz, and Mahesh P Thakur. 2019. Rice plant disease classification using transfer learning of deep convolution neural network. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences XLII-3/W6* (July 2019), 631–635. <https://doi.org/10.5194/isprs-archives-xlii-3-w6-631-2019>
- [6] Dogan, Y., & Keles, H. Y. (2020). Semi-supervised image attribute editing using generative adversarial networks. *Neurocomputing*, 401, 338-352.
- [7] Dogan, Y., & Keles, H. Y. (2022). Iterative facial image inpainting based on an encoder-generator architecture. *Neural Computing and Applications*, 34(12), 10001-10021.
- [8] ATAŞ, M., & DOĞAN, Y. (2015). Classification of Closed and Open Shell Pistachio Nuts by Machine Vision. In *International Conference on Advanced Technology Sciences*, Antalya.
- [9] ATAŞ, M., ÖZDEMİR, C., ATAŞ, İ., Ak, B., & ÖZEROĞLU, E. (2022). Biometric identification using panoramic dental radiographic images with few-shot learning. *Turkish Journal of Electrical Engineering and Computer Sciences*, 30(3), 1115-1126.
- [10] Dogan, Y. (2023). A new global pooling method for deep neural networks: Global average of top-k max-pooling. *Traitement du Signal*, Vol. 40, No. 2, pp. 577-587. <https://doi.org/10.18280/ts.400216>
- [11] Petchiammal A, Briskline Kiruba S, Murugan D, Pandarasamy Arjunan. (2022). Paddy Doctor: A Visual Image Dataset for Automated Paddy Disease Classification and Benchmarking. *IEEE Dataport*. <https://dx.doi.org/10.21227/hz4v-af08>
- [12] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.