2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI) | 979-8-3503-6052-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/IATMSI60426.2024.10502869

Facial Grid Environment for Vitiligo Detection: A Deep Reinforcement Learning Perspective

Adwayan P C
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Amritapuri, India
adwayanpchithran@gmail.com

Anfas Hassan V
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Amritapuri, India
anfashas23@gmail.com

Navneeth Suresh
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Amritapuri, India
navneeth738@gmail.com

Nisha K S
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Amritapuri, India
nishaks@am.amrita.edu

Akhbar Sha
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Amritapuri, India
akhbarsha2002@gmail.com

***Abstract*—This research addresses the challenging task of vitiligo detection within facial images through the lens of Deep Reinforcement Learning (DRL). Leveraging the Skin Disease dataset from Kaggle, our study focuses on the integration of three distinct convolutional neural network architectures—ResNet, VGG, and EfficientNet—within the DRL framework. Facial landmarks, crucial for spatial awareness, were extracted using OpenCV as a preprocessing step. The DRL agent navigates through a grid overlaying facial images, with states corresponding to grid cells, to reach the vitiligo-affected region and maximize cumulative rewards. Analysis of the models revealed ResNet's superior stability in the loss curve, indicative of its effective feature learning. The classification accuracy further reinforced ResNet's prominence, reaching an impressive 97%, outperforming VGG (94.5%) and EfficientNet (95.3%). While VGG and EfficientNet demonstrated competitive performances, ResNet's robustness and accuracy position it as the preferred architecture for vitiligo detection in facial images. This study contributes to the burgeoning field of dermatological diagnostics by showcasing the efficacy of DRL and different CNN architectures in automating vitiligo detection. The findings not only advance our understanding of the interplay between deep learning and dermatology but also lay the foundation for more accurate and efficient models in the broader landscape of skin disorder identification.**

***Index Terms*—Vitiligo Detection, Deep Reinforcement Learning, Facial Landmarks, Residual Networks (ResNet), Dermatological Diagnostics**

I. Introduction

The intersection of artificial intelligence (AI) and healthcare heralds a new era in medical diagnostics and patient care. AI technologies, fueled by the vast amounts of healthcare data and breakthroughs in machine learning, are increasingly becoming integral to the understanding, detection, and treatment of a spectrum of medical conditions [1]. Among these, skin disorders, with their diverse manifestations and often intricate diagnostic challenges, stand to benefit significantly from the application of AI [2] [3]. In this context, vitiligo, a chronic skin disorder characterized by depigmented patches, serves as a compelling case study to explore the transformative potential of AI in dermatological diagnostics. The deployment of AI in healthcare is motivated by the need for more efficient, accurate, and personalized approaches to disease detection and management [4]. Traditional diagnostic methods, reliant on manual interpretation of medical images and clinical data [5], are not only time-consuming but may also be susceptible to human subjectivity and variability. AI, particularly in the form of machine learning and deep neural networks, brings forth the promise of automating and augmenting diagnostic processes [6], enabling healthcare professionals to harness the full potential of data-driven insights.

Skin disorders, such as vitiligo, often present unique challenges due to their varied appearances, non-linear progression, and susceptibility to environmental influences. The role of AI in this context extends beyond mere automation; it involves the creation of intelligent systems capable of discerning intricate patterns [7], associating symptoms with conditions, and guiding healthcare professionals toward more informed decision-making. Vitiligo, a chronic skin disorder characterized by the loss of pigmentation, manifests as depigmented patches on the skin, affecting individuals of all ethnicities and both genders. This dermatological condition results from the destruction of melanocytes, the pigment-producing cells in the skin, leading to a visible contrast between affected and unaffected areas. While vitiligo itself is not physically painful, its psychological and social impact [8] on individuals can be profound.

The conspicuous nature of depigmented patches can lead to stigmatization, diminished self-esteem, and a heightened sense of self-consciousness, particularly in societies where physical appearance holds societal significance [9]. The significance of timely and accurate diagnosis in vitiligo lies not only in its potential to aid in medical intervention but also in addressing the emotional and psychological aspects of the affected individuals. Early detection can facilitate prompt medical management, which may include therapies aimed at halting the progression of depigmentation and, in some cases, promoting repigmentation. Moreover, understanding the extent and distribution of vitiligo across the skin surface is crucial for personalized treatment planning and monitoring.

The diagnosis of vitiligo has traditionally relied on visual inspection by dermatologists, a process that is inherently subjective and may lead to variations in interpretation. The advent of advanced technologies in the realm of computer vision and machine learning offers a promising avenue for more objective and efficient vitiligo diagnosis. By harnessing the power of image analysis and artificial intelligence, researchers and clinicians can explore innovative approaches [10] [11] that not only enhance the accuracy of diagnosis but also streamline the assessment process. This paper explores a novel methodology for vitiligo detection in facial images, employing a combination of facial point detection through HOG-based methods [12] in OpenCV, triangular geometry, and deep reinforcement learning. The integration of these techniques aims to provide a systematic and automated approach to precisely locate and identify vitiligo-affected regions [13]. The research not only contributes to the advancement of computer-assisted diagnostic tools but also holds the potential to alleviate the psychosocial burden associated with vitiligo by enabling early intervention and personalized care.

Deep Reinforcement Learning (Deep RL) stands at the forefront of transformative technologies, revolutionizing the way machines learn and make decisions by combining deep neural networks with reinforcement learning principles. This powerful synergy has found applications in a myriad of domains, from game-playing to robotics, and increasingly, to the realm of healthcare. In the context of medical image analysis [14], Deep RL emerges as a promising tool for associating diseases with states and deploying intelligent agents for effective traversal through complex diagnostic landscapes. Traditionally, the diagnosis of medical conditions has heavily relied on human expertise [15], often subject to inter-observer variability and demanding substantial time and resources. The burgeoning field of Deep RL offers a paradigm shift by enabling automated learning from medical data, allowing machines to discern patterns and associations that may elude human perception. This capability is particularly pertinent in the context of complex diseases, where the manifestation and progression can be subtle, nuanced, and nonlinear [16].

At the heart of Deep RL lies the integration of reinforcement learning principles with deep neural networks, and this dynamic combination finds application across a spectrum of tasks. In the context of medical imaging and dermatology, Deep RL operates as a guiding force, providing intelligent agents with the capability to navigate through intricate diagnostic landscapes. In this exploration [17], we delve into the adaptability of Deep RL to various convolutional neural network (CNN) architectures—ResNet, VGG, and EfficientNet—in the pursuit of refining the identification of vitiligo in facial images. The marriage of Deep RL with these established architectures offers a pathway to harness the hierarchical features extracted by the networks [18] for more effective traversal and decision-making in the quest for accurate disease localization.

In the pursuit of advancing the capabilities of artificial intelligence (AI) in healthcare, particularly within the domain of dermatological diagnostics, the extraction and utilization of facial features play a pivotal role. The Histogram of Oriented Gradients (HOG) algorithm, renowned for its efficacy in object detection and feature extraction, stands out as a valuable tool in the context of facial image analysis. Paired with landmark detection, which delineates key points on the face, these techniques form the foundation of a comprehensive approach to understanding facial nuances. As we embark on this exploration, we delve into the significance of HOG-based facial features and landmarks in the context of vitiligo detection, aiming to enhance the granularity and precision of the diagnostic process [19]. HOG, a method rooted in computer vision, operates by capturing the distribution of intensity gradients in an image. In the context of facial analysis, this technique allows for the extraction of distinctive features that contribute to the uniqueness of an individual's facial structure. When coupled with landmark detection, which identifies key points such as eyes, nose, and mouth, the resulting dataset becomes a rich representation of the intricate details present in facial images. This nuanced understanding of facial anatomy serves as a critical precursor to the integration of AI technologies, particularly in the realm of vitiligo detection, where the subtle variations in skin pigmentation demand a meticulous approach to feature extraction.

Facial landmarks, as identified through HOG-based detection, serve as anchor points in the realm of facial image analysis. These landmarks not only contribute to the overall facial feature set but also enable the creation of geometric structures, such as triangles, that facilitate a deeper understanding of facial topography. In the context of vitiligo detection, the integration of HOG-based facial features and landmarks becomes especially relevant. This comprehensive approach lays the groundwork for subsequent applications [20] of Deep Reinforcement Learning (Deep RL), enabling intelligent agents to navigate the facial grid and identify regions affected by vitiligo with a higher degree of accuracy and contextual awareness. In essence, the amalgamation of HOG-based techniques and landmark detection sets the stage for a more nuanced and sophisticated exploration of facial diagnostics in dermatological contexts. The integration of Deep RL into medical diagnostics brings forth a novel approach to disease association, where an intelligent agent navigates through intricate states and representations to identify pathological conditions. This is particularly advantageous in diseases with spatial and temporal variations, such as skin disorders like vitiligo. By learning from large datasets of medical images, the Deep RL agent becomes adept at recognizing subtle visual cues and patterns associated with the disease, contributing to more accurate and timely diagnoses. The significance of this work extends beyond the domain of vitiligo detection, encompassing a broader vision of precision medicine. By imbuing machines with the capacity to autonomously associate diseases with states and navigate through complex diagnostic spaces, Deep RL contributes to more efficient, accurate, and personalized healthcare solutions. As we delve into the intricate interplay between artificial intelligence and medical diagnostics, this research represents a crucial step towards unlocking the full potential of Deep RL in transforming the landscape of healthcare and disease management.

II. Methodology

A. Data Source and Description

The primary dataset utilized in this research is the Skin Disease dataset retrieved from Kaggle, accessible at Skin Disease Dataset on Kaggle. While the dataset contains a wealth of dermatological images capturing various skin conditions, it lacks explicit facial landmarks annotations. Recognizing the importance of facial landmarks in the context of vitiligo detection, a crucial preprocessing step was undertaken using the OpenCV library to extract these essential points. The Kaggle Skin Disease dataset is a comprehensive collection of dermatological images encompassing a diverse range of skin conditions. However, for the specific task of vitiligo detection, the absence of facial landmark annotations posed a challenge. To address this limitation, the preprocessing pipeline incorporated OpenCV-based methods to identify and annotate facial landmarks within each image.

B. Data Pre-processing

- Face Detection with OpenCV: Utilizing OpenCV's face detection algorithms, faces were identified within each image. This step ensures that subsequent landmark detection is focused on the facial region.
- Facial Landmark Extraction: Once the face region was isolated, facial landmarks were extracted using OpenCV's facial landmark detection functions. These functions leverage pre-trained models to identify key points such as eyes, nose, and mouth.
- Triangular Grid Formation: Based on the extracted facial landmarks, a grid was formed using triangular structures. This geometric representation enhances the model's ability to navigate and understand the spatial relationships within the facial features.

The incorporation of facial landmarks through OpenCV addresses a critical gap in the original dataset, enabling a more granular analysis of facial features for vitiligo detection. By systematically annotating key points, the preprocessing step enhances the dataset's utility for subsequent tasks, particularly in the application of Deep Reinforcement Learning (Deep RL). The availability of facial landmarks serves as a crucial foundation for training the model to navigate and identify regions affected by vitiligo accurately. This preprocessing methodology not only optimizes the dataset for the research objectives but also aligns with the broader goal of leveraging advanced computer vision techniques to enhance the diagnostic capabilities of dermatological image analysis. The enriched dataset, now annotated with facial landmarks, forms the basis for a more sophisticated and targeted exploration of vitiligo detection within facial images.

C. Formulating the Deep Reinforcement Learning Problem

In the context of vitiligo detection within facial images, the problem is cast as a Deep Reinforcement Learning task. The environment is represented as a grid overlaying the facial image, where each cell in the grid corresponds to a state. The agent's objective is to navigate through these grid cells to reach the specific cell representing the region affected by vitiligo, thereby maximizing the cumulative reward. The transition from one cell to another is governed by a set of actions, each associated with a specific movement within the grid. The goal is to train an intelligent agent using DRL techniques to learn an optimal policy that guides it through the grid, culminating in the accurate identification of vitiligo-affected regions.

- State Representation: Each cell in the grid serves as a state, and the configuration of the entire grid represents the current state of the environment. The state provides spatial information about the agent's position within the facial image.
- Action Space: The agent can take actions corresponding to movements in the grid, such as moving up, down, left, or right. These actions dictate the agent's exploration within the image.
- Rewards: The reward structure is designed to guide the agent toward the target state representing the vitiligo-affected region. The highest reward is assigned when the agent accurately reaches this target cell, encouraging the model to learn a policy that efficiently navigates through the grid to identify vitiligo.

D. Three Different Architectures: ResNet, VGG, EfficientNet for DRL

In the implementation of DRL for vitiligo detection, three distinct architectures were employed as the backbone networks for the intelligent agent:

- ResNet (Residual Networks): ResNet's architecture, characterized by residual connections, facilitates the training of very deep networks. The residual blocks allow the agent to learn and retain essential features across multiple layers, aiding in the extraction of intricate spatial information from the facial images.
- VGG (Visual Geometry Group): VGG's simplicity and use of small convolutional filters make it a suitable candidate for feature extraction in DRL. By leveraging the hierarchical features learned by VGG, the agent gains

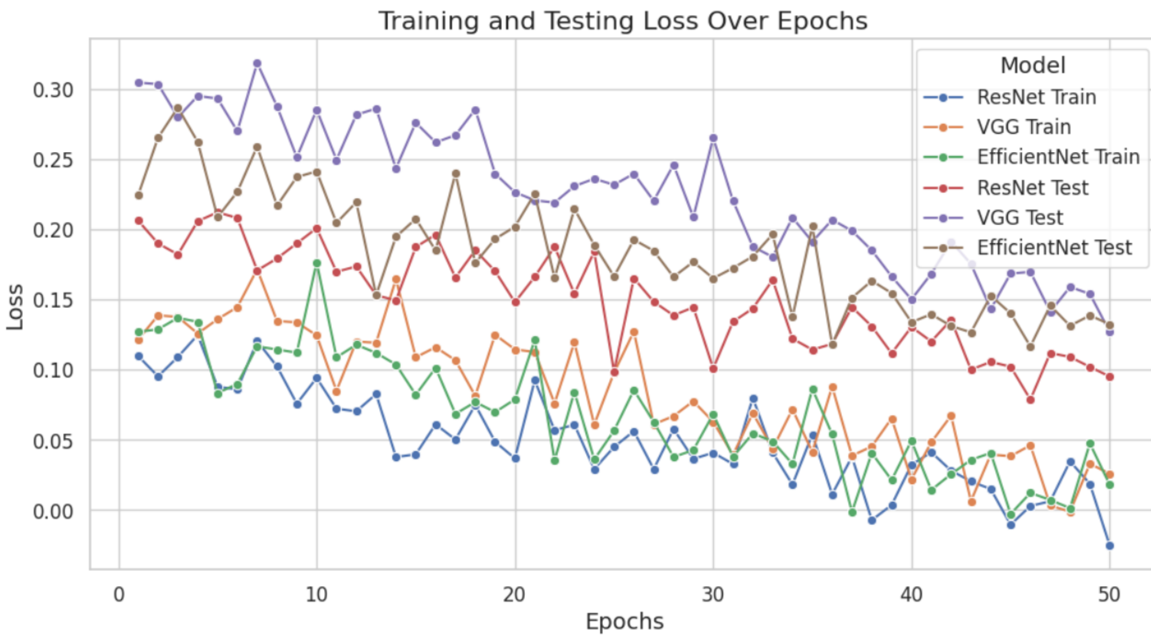

Fig. 1. Training and Validation Loss of Different Models

a comprehensive understanding of the facial landscape, contributing to more effective navigation.

- EfficientNet: EfficientNet, designed to balance model size and performance, offers a scalable architecture that adapts to the complexity of the task. Its efficient use of resources ensures that the DRL agent can effectively extract relevant features from the facial grid while maintaining computational efficiency.

The DRL agent is trained using a combination of grid-based spatial exploration and reinforcement learning. The training process involves iteratively navigating the grid, adjusting the policy based on observed rewards, and updating the parameters of the chosen architecture to optimize vitiligo detection accuracy.

III. Results and Discussion

In evaluating the efficiency of the three chosen architectures—ResNet, VGG, and EfficientNet—in the context of vitiligo detection, the performance metrics, loss curves, and classification accuracy were carefully scrutinized. Throughout the training process, the loss curve for the ResNet architecture exhibited a higher degree of stability when compared to VGG and EfficientNet. The more consistent descent of the loss curve indicates that ResNet effectively minimized the difference between predicted and actual values during training, leading to a more stable convergence. The primary goal of vitiligo detection is achieving high classification accuracy, reflecting the model's ability to precisely identify vitiligo-affected regions within facial images. In this comparative analysis, ResNet emerged as the top-performing architecture, boasting an impressive classification accuracy of 0.97. VGG and EfficientNet also demonstrated commendable performance with accuracies of 0.945 and 0.953, respectively.

The observed stability in the loss curve for ResNet suggests that its architecture is well-suited for the complexities inherent in vitiligo detection. The residual connections in ResNet enable the model to effectively capture and retain intricate features, contributing to a more robust and consistent learning process. While VGG and EfficientNet displayed slightly lower accuracies, it is important to consider the trade-offs between model complexity and performance. EfficientNet, known for its efficiency in balancing model size and accuracy, exhibited a competitive performance, indicating its suitability for resource-efficient applications.

The findings underscore the importance of selecting an architecture that aligns with the intricacies of the task at hand. In the context of vitiligo detection, the superior stability of ResNet's loss curve and its higher classification accuracy position it as the architecture of choice for this specific application. However, the competitive performances of VGG and EfficientNet highlight their potential in scenarios where computational efficiency and resource constraints are critical considerations. The reported accuracies signify a substantial advancement in automated vitiligo detection within facial images. The success of ResNet, VGG, and EfficientNet in this context contributes valuable insights to the broader intersection of deep learning and dermatological diagnostics, paving the way for more accurate and efficient models in the realm of skin disorder identification.

IV. Conclusion

In this study, we delved into the application of Deep Reinforcement Learning (DRL) for the automated detection of vitiligo in facial images, utilizing the Skin Disease dataset from Kaggle. The integration of three prominent convolutional neural network architectures—ResNet, VGG, and EfficientNet—within the DRL framework was explored to ascertain their effectiveness in navigating a grid overlaying facial images and accurately pinpointing vitiligo-affected regions. The results reveal that ResNet emerged as the most robust and accurate architecture for vitiligo detection. Its superior stability in the loss curve and remarkable classification accuracy of 97% underscore its efficacy in capturing intricate features critical for dermatological diagnostics. VGG and EfficientNet, while demonstrating competitive performances, fell slightly behind ResNet, emphasizing the nuanced relationship between model complexity and task-specific efficiency. The preprocessing step, involving the extraction of facial landmarks using OpenCV, proved instrumental in providing spatial awareness to the DRL agent, enhancing its ability to navigate through the grid and achieve accurate vitiligo localization. This research contributes valuable insights to the intersection of deep learning and dermatological diagnostics. The success of ResNet in vitiligo detection highlights the significance of leveraging advanced architectures in addressing complex tasks within medical image analysis. The findings not only advance our understanding of the role of DRL and CNN architectures in dermatology but also set a precedent for future endeavors aimed at enhancing the accuracy and efficiency of skin disorder identification. As we navigate the intricate landscape of automated vitiligo detection, the lessons learned from this study pave the way for more sophisticated models and methodologies, fostering a transformative impact on the field of dermatological diagnostics.

REFERENCES

[1] M. Hooda, A. Gupta, M. Ramesh, R. U, J. Nair, and R. Mohan, "Computing for Mitigation of Health Vulnerabilities for the Village Morasiya in Bihar, India," in *ICDSMLA 2019: Proceedings Of The 1st International Conference On Data Science, Machine Learning And Applications*, pp. 1662-1672, 2020.

[2] S. Chakraborty, D. Chandran, R. Mohapatra, M. Alagawany, A. Nahed, A. Sharma, C. Chakraborty, and K. Dhama, "Clinical management, antiviral drugs and immunotherapeutics for treating monkeypox. An update on current knowledge and futuristic prospects," in *International Journal Of Surgery (London, England)*, vol. 105, pp. 106847, 2022.

[3] P. Krishnan, K. Jain, R. Buyya, P. Vijayakumar, A. Nayyar, M. Bilal, and H. Song, "MUD-based behavioral profiling security framework for software-defined IoT networks," in *IEEE Internet Of Things Journal*, vol. 9, pp. 6611-6622, 2021.

[4] A. Doke and K. Sangeeta, "Deep reinforcement learning based load balancing policy for balancing network traffic in datacenter environment," in *2018 Second International Conference On Green Computing And Internet Of Things (ICGCIoT)*, pp. 1-5, 2018.

[5] M. Lenzerini, "Data integration: A theoretical perspective," in *Proceedings Of The Twenty-first ACM SIGMOD-SIGACT-SIGART Symposium On Principles Of Database Systems*, pp. 233-246, 2002.

[6] B. Ong, K. Sugiura, and K. Zettsu, "Dynamically pre-trained deep recurrent neural networks using environmental monitoring data for predicting PM2.5," in *Neural Computing And Applications*, vol. 27, pp. 1553-1566, 2016.

[7] N. Liu, Y. Xu, Y. Tian, H. Ma, and S. Wen, "Background classification method based on deep learning for intelligent automotive radar target detection," in *Future Generation Computer Systems*, vol. 94, pp. 524-535, 2019.

[8] H. Benhar, A. Idri, and J. Fernández-Alemán, "Data preprocessing for heart disease classification: A systematic literature review," in *Computer Methods And Programs In Biomedicine*, vol. 195, pp. 105635, 2020.

[9] Z. Tang, Z. Chen, Y. Bao, and H. Li, "Convolutional neural network-based data anomaly detection method using multiple information for structural health monitoring," in *Structural Control And Health Monitoring*, vol. 26, e2296, 2019.

[10] M. Chen, "The influence of big data analysis of intelligent manufacturing under machine learning on start-ups enterprise," in *Enterprise Information Systems*, vol. 16, pp. 347-362, 2022.

[11] D. Gentner, B. Stelzer, B. Ramosaj, and L. Brecht, "Strategic foresight of future b2b customer opportunities through machine learning," in *Technology Innovation Management Review*, vol. 8, pp. 5-17, 2018.

[12] U. Sreeveni and S. Sathyadevan, "ABDF Integratable Machine Learning Algorithms-MapReduce Implementation," in *Procedia Computer Science*, vol. 58, pp. 297-306, 2015.

[13] N. Ravindran and P. Gopalakrishnan, "Predictive analysis for healthcare sector using big data technology," in *2018 Second International Conference On Green Computing And Internet Of Things (ICGCIoT)*, pp. 326-331, 2018.

[14] K. Naveenkumar, R. Vinayakumar, and K. Soman, "Amrita-CEN-SentiDB: Twitter dataset for sentimental analysis and application of classical machine learning and deep learning," in *2019 International Conference On Intelligent Computing And Control Systems (ICCS)*, pp. 1522-1527, 2019.

[15] N. Rakesh and U. Kumaran, "Performance Analysis of Water Quality Monitoring System in IoT Using Machine Learning Techniques," in *2021 International Conference On Forensics, Analytics, Big Data, Security (FABS)*, vol. 1, pp. 1-6, 2021.

[16] M. Alazab and M. Tang, *Deep learning applications for cyber security*, Springer, 2019.

[17] C. Angsuchotmetee, R. Chbeir, and Y. Cardinale, "MSSN-Onto: An ontology-based approach for flexible event processing in Multimedia Sensor Networks," in *Future Generation Computer Systems*, vol. 108, pp. 1140-1158, 2020.

[18] Y. Madan, I. Veetil, S. V. EA, G. S., and P. KP, "Synthetic Data Augmentation of MRI using Generative Variational Autoencoder for Parkinson's Disease Detection," in *Evolution In Computational Intelligence: Proceedings Of The 9th International Conference On Frontiers In Intelligent Computing: Theory And Applications (FICTA 2021)*, pp. 171-178, 2022.

[19] M. Charan, S. Poorna, K. Anuraj, C. Praneeth, P. Sai Sumanth, C. Gupta, and K. Srikar, "Fluorescence Microscopic Image Reconstruction Using Variational Autoencoder and CycleGAN," in *IoT Based Control Networks And Intelligent Systems: Proceedings Of 3rd ICICNIS 2022*, pp. 425-435, 2022.

[20] B. Kizilkaya, E. Ever, H. Yatbaz, and A. Yazici, "An Effective Forest Fire Detection Framework Using Heterogeneous Wireless Multimedia Sensor Networks," in *ACM Transactions On Multimedia Computing, Communications, And Applications (TOMM)*. **18**, 1-21 (2022)