Comparative Evaluation of Machine Learning Techniques for Classifying Multi-Label Skin Cancer

[1] Manjusha Nambair, [2] A. Vanaja,

[1]Assistant Professor, Megha Institute of Engineering & Technology for Women, Ghatkesar.
[2] MCA Student, Megha Institute of Engineering & Technology for Women, Ghatkesar.

Abstract

Despite the frequency of skin cancer, many people may not know the symptoms or what they can do to protect themselves. Skin cancer is the fourth most frequent illness globally, and its incidence and fatality rate are rising. Therefore, it is crucial to detect cancer in its early stages in order to stop it from spreading. Finding and categorizing skin tumors using various labels is the goal of this study, which employs state-of-the-art techniques in machine learning and image processing. While preprocessing techniques aid in removing unnecessary and irrelevant features from the label encoder, standard features assist in standardizing the range of functioning by scaling the input variance unit. All classifiers, employing a variety of ML techniques, were evaluated on the HAM10000 metadata dataset. For this experimental study, we utilized the HAM10000 metadata dataset, which contains information on seven different kinds of skin cancer. Data analysis revealed that SVM, DT, and GNB were the top classifiers among the machine learning algorithms.

Keywords:

Skin cancer, SVM, DT, GNB, classification, machine learning, and multi-class

INTRODUCTION

Cancer of the skin affects a wide variety of animals and people in various parts of the world. However, at this moment it provides a clear advantage. The increasing incidence of skin cancer is causing immense suffering for individuals around the world. Skin cancer is the fourth most common cancer worldwide. The elderly and young are most at risk, yet everyone may be affected at any moment [1]. The possibility of a surgical operation curing the ailment depends on how quickly it is identified. Melanoma, basal, and squamous cells are only a few of the many types [2]. Melanoma is one of the malignancies that might be unexpected. Damage occurs to skin cells, hair follicles, and mucosal membranes. It is possible for skin cancer to develop from almost any damaged skin cell. This disorder, which affects almost everyone, arises when healthy skin cells change into aberrant, malignant ones. Skin malignancies, known as carcinomas, may develop in any part of the body. High quantities of ultraviolet radiation from the sun have been linked to skin cancer, according to studies [3]. Skin cancer affects a lot of people all over the world, but not everyone is aware of it. The sun's ultraviolet (UV) rays are to blame for the skin cancer that they cause. The atmosphere is a pathway for the sun's radiation, which includes ultraviolet (UV) rays. It might show up anywhere on your body and in a variety of ways. While it may be impossible to completely eradicate unclear skin cancer cells, there are strategies to lessen their effect. One serious problem that may lead to serious harm is skin cancer [4]. Avoid cancer waves by being aware of your legal obligations and taking the necessary safeguards. A substantial percentage of people in the US are sick [5]. In 2012, there were over 63,000 new cases of melanoma, out of the millions of new occurrences of non-melanoma skin cancers (NMSC), according to the Skin Cancer Foundation. Because of this, melanoma is the skin cancer that kills the most people. The uncontrolled proliferation of skin cells is the root cause of skin cancer. This kind of skin cancer is known as non-melanoma skin cancer (NMSC) when it occurs on the skin's surface. Melanoma, in contrast, occurs when skin cells proliferate farther into the dermis [6].

II. RELATED WORK

Over the last two decades, researchers in academia and industry have created machine learning methods to automate the process of skin cancer classification. The worldwide suffering caused by skin cancer is increasing, and Nazia Hameed et al. suggested a way to improve approaches to many disorders by using deep learning and machine learning. Classification algorithms can determine the best ways to improve different skin lesions [7]. In order to analyze the effects of skin color purification and segmentation on human disorders, A. Murugan et al. used anecdotal photographs to regulate mean feature extraction. When used together, the SVM and RF algorithms outperform competing methods [8]. In their presentation, Carolina Magalhaes et al. discussed how infrared thermography and machine learning may be used to identify skin cancer. One example of how confusion matrix simplification may increase prediction accuracy is the experimental skin cancer detection that used ensemble learning based on input thermal characteristics [9]. Melanoma develops, as proposed by Mehwish Dildar et al., when injured skin cells continue to transmit DNA to other organs. When it comes to early cancer detection and treatment, factors including color, form, symmetry, and symptom strength should be taken into account. Results might be better when using machine learning and deep learning algorithms to identify skin cancer, despite the fact that numerous research have done so [10]. By analyzing the polarization deep learning illustration image and illness statistics evaluation pattern, Yuheng Wang et al. evaluated a cancer detection supplement that aims to improve our understanding of human cancer. However, when comparing deep learning and machine learning algorithms, two types of skin cancer were identified: benign lesions and malignant lesions [11]. Rashmi Patil et al. found that although machine learning algorithms might bother people, they are able to detect melanoma, the most serious and concerning kind of cancer. For the purpose of quantifying the loss function and resolving the text processing melanoma tumor thickness classification problem, the suggested method depended on the CNN model [12]. According to Ravi Manne et al., physicians have shown that deep learning techniques called convolution neural networks may help distinguish between skin malignancies. The analyzed convolutional neural network (CNN) model used deep learning approaches to increase accuracy by minimizing image misclassifications [13].

III. METHODOLOGY

Machine learning techniques as SGD Classifier, Logistic Regression, KNN, GNB, and DT are used in the proposed method for multilabel data classification. The typical workflow consists of three phases: preprocessing, classification, and performance evaluation. Figure 1 depicts the suggested system design.

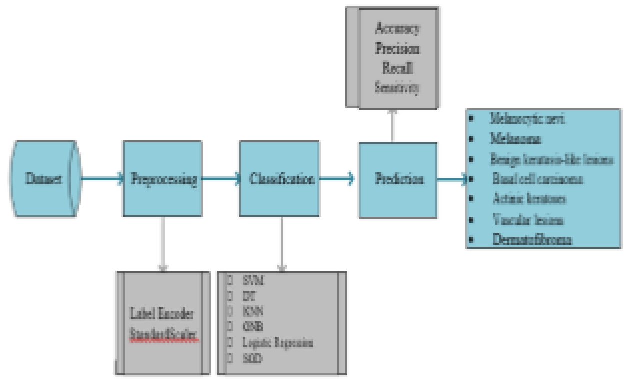


Fig 1: Proposed methodology to Skin cancer classification

A. Getting Ready If we want machine learning technologies to work as efficiently and accurately as possible, we need clean data first. Each attribute's coefficient is calculated using a different approach. In order for the machine learning algorithm to understand the label data, the label encoder characteristics were used to transform it into a numerical representation. To further institutionalize the spectrum of processes in the input data, a common scalar component was scaled to unit variance and reused [13]. Section B: Arranging With multi-label data, many classification methods were used, such as DT, KNN, GNB, Logistic Regression, SGD Classifier, and Logistic Regression. Applying a variety of labels to categorize SVM Using SVM, a supervised approach, is the way to go for regression and classification challenges. Problems involving binary classification are also addressed by the SVM using the same approach. Substantial problems are included within the multi-classification problem. The One versus. All (OVA) technique stands out as a remarkable way for multi-classification on the problem statement. One versus. All separates the classes further once the hyperplane splits them, creating two sets of classes for each class point and another set for all other points. The Greenline maximizes the distance between the green point and all other points, as seen in Figure 2 [14].

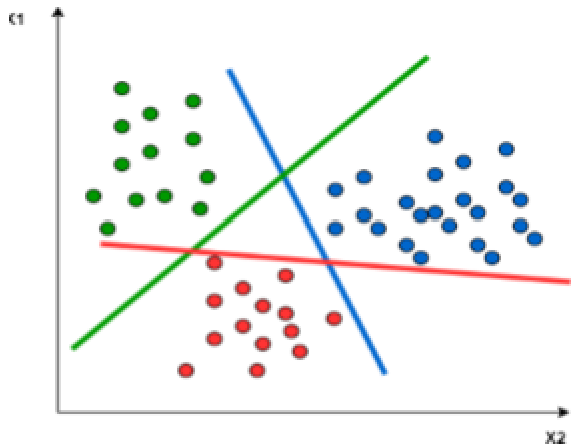
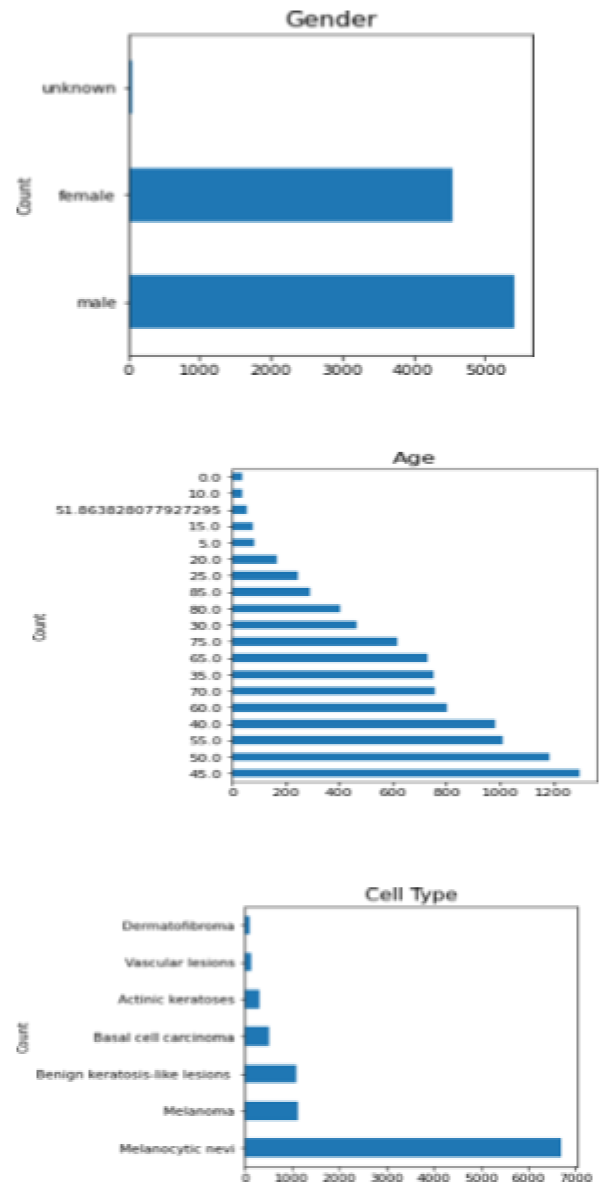


Figure 2: One vs All approach

Applying DT to multiple-label classification A methodical strategy for multi-label classification is the DT classifier. It shapes itself like a tree and bases its queries on characteristics and attributes. Each internal root node organizes the data into different records based on a range of parameters. The many nodes on the tree represent different types of data [15]. Applying a variety of labels to categorize KNN A supervised machine learning method for classification is Karl-Nielsen-Nielsen (KNN). The KNN method is insensitive to data structure. You may use the geometric distance formula to find out how far apart the two feature vectors are [16]. Applying a variety of labels to categorize GNB Among the many forms of probabilistic machine learning, Neural Networks (NB) use Bye's theorem for data classification. In order to determine whether the training data follows a normal or gaussian distribution, we use several functions. Finding the mean and variance of X is the first step in GNB, followed by plugging in the probability thickness of the normal distribution [17]. Applying a variety of labels to categorize Logistic regression model When it comes to classification, supervised machine learning methods like the LR algorithm are useful. As shown by one-vs-rest [18], the LR method may solve problems associated with multiclass classification, such modifying the loss functions and predicting the distribution to a multinomial probability. Applying a variety of labels to categorize Convex loss functions include logistic regression and (linear) support vector machines; SGD offers an easy and fast way to get the right linear classifiers and regressors. Despite SGD's lengthy history in the machine learning community, its rise to prominence in the big data arena is very recent [19]. Part C: Dataset The HAM10000 metadata data set was taken from the Kaggle repository and used for this research. Data is gathered from many sources and then stored in various ways. Dataset is an open-source machine learning database that researchers may use to study how machine learning tasks handle large amounts of data. With the use of machine learning models, it could track the dataset's expected features in real time across two sets of data: training and testing [20]. Figure 3 shows the findings' partial distributions for location, cell type, age, and gender. It is evident from the data's behavior that applying machine learning algorithms to chaotic data will provide undesirable outcomes. While 80% of the dataset was reused for training, only 20% was used for testing.



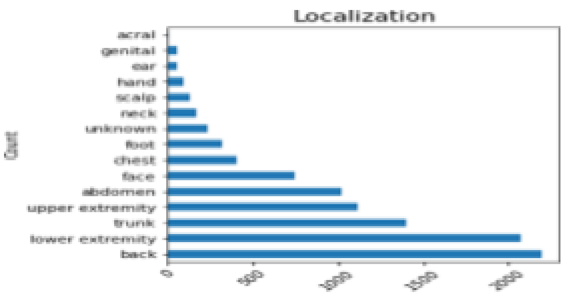


Fig 3: Distribution of different attributes values from the dataset

IV. RESULT AND DISCUSSION

A variety of classification models and their analyses are presented in this section. SVM, DT, KNN, GNB, Logistic Regression, and SGD are the six machine learning techniques that were applied on the HAM10000_metadata dataset in order to evaluate each classifier's performance. Prior to classifiers, the data is first standardized and normalized. In this experiment, the data is trained and evaluated on various machine learning classifiers. 20% of the data was used for testing, while the remaining 80% was used for training. Tables 1 through 6 display the outcomes of various machine learning methods. Below is a detailed explanation of label mapping:

- 0: Melanocytic Nevi (NV)
- 1: Melanoma (MEL)
- 2: Benign Keratosis-Like Lesions (BKL)
- 3: Basal Cell Carcinoma (BCC)
- 4: Intraepithelial Carcinoma / Bowen's Disease (AKIEC)
- 5: Vascular Lesions (VASC)
- 6: Dermatofibroma (DF)

Table 1: Multi labels classification SVM

	precision	recall	f1-score	support
0	1.00	1.00	1.00	61
1	1.00	1.00	1.00	96
2	1.00	1.00	1.00	228
3	1.00	1.00	1.00	37
4	1.00	1.00	1.00	1327
5	1.00	1.00	1.00	222
6	1.00	1.00	1.00	32
accuracy			1.00	2003
macro avg	1.00	1.00	1.00	2003
weighted avg	1.00	1.00	1.00	2003

Table 2: Multi labels classification DT

	precision	recall	f1-score	support
0	1.00	1.00	1.00	61
1	1.00	1.00	1.00	96
2	1.00	1.00	1.00	228
3	1.00	1.00	1.00	37
4	1.00	1.00	1.00	1327
5	1.00	1.00	1.00	222
6	1.00	1.00	1.00	32
accuracy			1.00	2003
macro avg	1.00	1.00	1.00	2003
weighted avg	1.00	1.00	1.00	2003

Table 3: Multi labels classification KNN

	precision	recall	f1-score	support
0	0.85	0.82	0.83	61
1	0.86	0.85	0.86	96
2	0.91	0.99	0.95	228
3	0.94	0.43	0.59	37
4	0.96	0.99	0.98	1327
5	0.95	0.89	0.92	222
6	1.00	0.06	0.12	32
accuracy			0.95	2003
macro avg	0.92	0.72	0.75	2003
weighted avg	0.95	0.95	0.94	2003

Table 4: Multi labels classification GNB

	precision	recall	f1-score	support
0	1.00	1.00	1.00	61
1	1.00	1.00	1.00	96
2	1.00	1.00	1.00	228
3	1.00	1.00	1.00	37
4	1.00	1.00	1.00	1327
5	1.00	1.00	1.00	222
6	1.00	1.00	1.00	32
accuracy			1.00	2003
macro avg	1.00	1.00	1.00	2003
weighted avg	1.00	1.00	1.00	2003

Table 5: Multi labels classification Logistic Regression

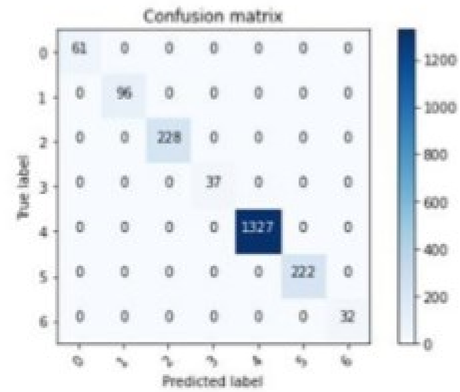
	precision	recall	f1-score	support
0	1.00	1.00	1.00	61
1	1.00	1.00	1.00	96
2	0.97	1.00	0.98	228
3	1.00	0.38	0.55	37
4	1.00	1.00	1.00	1327
5	0.93	1.00	0.97	222
6	1.00	1.00	1.00	32
accuracy			0.99	2003
macro avg	0.99	0.91	0.93	2003
weighted avg	0.99	0.99	0.99	2003

Table 6: Multi labels classification SGD Classifier

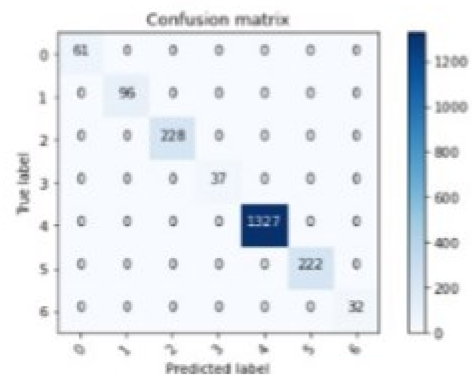
	precision	recall	f1-score	support
0	1.00	1.00	1.00	61
1	0.00	0.00	0.00	96
2	0.69	0.36	0.47	228
3	0.00	0.00	0.00	37
4	0.93	0.98	0.96	1327
5	0.43	0.74	0.54	222
6	0.00	0.00	0.00	32
accuracy			0.80	2003
macro avg	0.44	0.44	0.42	2003
weighted avg	0.77	0.80	0.78	2003

A. Matrix of Confusion The confusion matrix for skin cancer classification into seven distinct classes and prediction outcomes is shown in Figure 4. Confusion matrices make it simple to visualize the results of machine learning systems. Additionally, imbalanced data and more irrelevant instances of the data related to other classes are present in the majority of datasets. By using confusion matrix techniques, the model can predict every feature in order to obtain the greatest accuracy score. The confusion matrix demonstrates that throughout training and testing, the SVM, DT, and GNB performed better, achieving 100% accuracy, precision, recall, and F1-score.

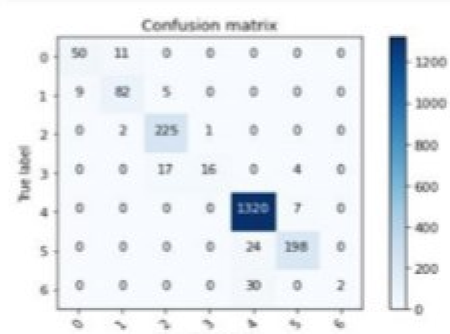
SVM



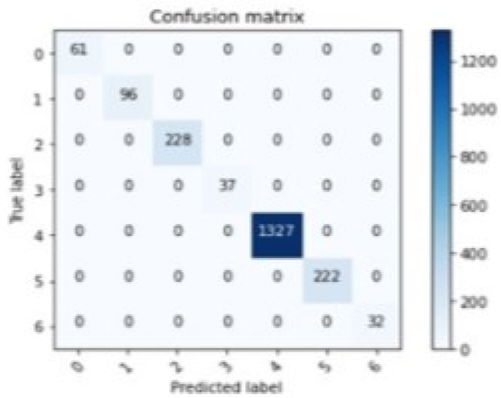
DT



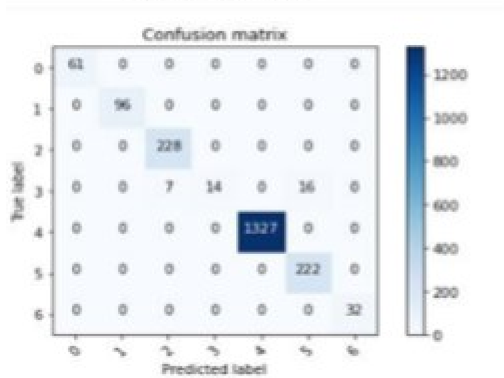
KNN



GNB



Logistic Regression



SGD Classifier

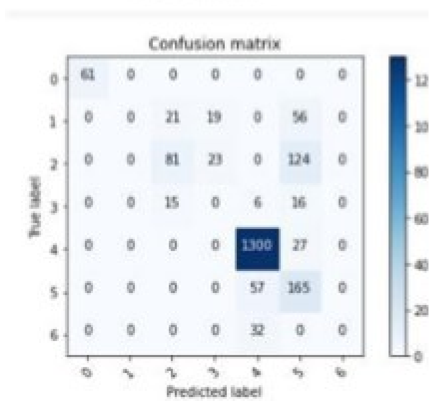


Figure 4: Confusion Matrix of different classifier (SVM, DT, KNN, GNB, Logistic Regression, SGD)

B. Total Outcome According to Table 7 and Figure 5, the results of the examination of all classifiers seem to indicate that the SVM, DT, and GNB have improved their accuracy the most.

Table 7: Comparison of all ML classifier

Machine Learning Classifier	Mean of Machine Learning Classifiers				
	Acc	Prec	Rec	F1-sco	Support
SVM	100%	100%	100%	100%	2003
DT	100%	100%	100%	100%	2003
KNN	95%	92%	72%	75%	2003
GNB	100%	100%	100%	100%	2003
LR	99%	99%	91%	93%	2003
SGD	80%	44%	44%	42%	2003

With 100% accuracy, 100% precision, 100% recall, and 100% F1Score, the SVM, DT, and GNB demonstrate strong performance. 91% recall, 93% F1-score, 99% accuracy, and 99% precision have all been attained via the logistic regression. 95% accuracy, 92% precision, 72% recall, and a 75% F1-score were attained with the KNN classifier. The SGD classifier has obtained the minimal performance accuracy of 80% accuracy, 44% precision, 44% recall, and 42% F1-score. The support vector machine, decision tree, and gaussian naive bayes classifiers are the best-performing algorithms among the aforementioned machine learning techniques. Table 7 and Figure 5 show that these classifiers had the highest accuracy.

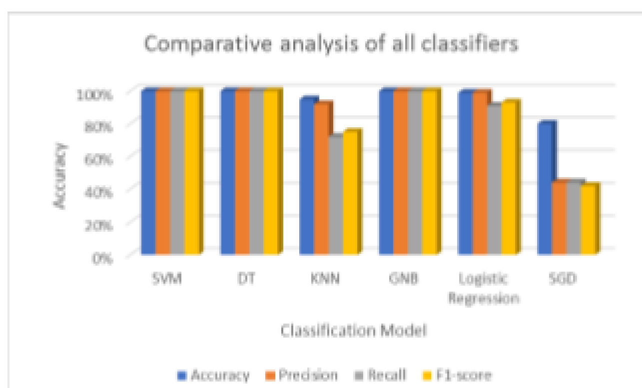


Figure 5: Comparison of all classifiers

V. CONCLUSION

Skin cancer is a leading cause of death annually. The annual diagnostic rate for melanoma is 53.3%, and there are 5.4 million new cases recorded globally. Starting with the skin cancer HAM10000_metadata dataset, we compared seven distinct label datasets using diverging machine learning algorithms. Classifiers with the best performance (100 percent accuracy) have been identified using several machine learning methods like SVM, DT, and GNB. SGD attained 80% accuracy, KNN 95% accuracy, and logistic regression 99% accuracy. We want to apply more categorization algorithms as we go along to gauge the dataset's efficacy.

REFERENCES

[1] A. Salam, F. Ullah, M. Imad, and M. A. Hassan, "Diagnosing of Dermoscopic Images using Machine Learning approaches for Melanoma Detection," in 2020 IEEE 23rd International Multitopic Conference (INMIC), 2020: IEEE, pp. 1-5.

[2] K. Das et al., "Machine Learning and Its Application in Skin Cancer", International Journal of Environmental Research and Public Health, vol. 18, no. 24, p. 13409, 2021. Available: 10.3390/ijerph182413409.

[3] E. Jana, R. Subban and S. Saraswathi, "Research on Skin Cancer Cell Detection Using Image Processing", 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2017. Available: 10.1109/iccic.2017.8524554.

[4] P. Dubal, S. Bhatt, C. Joglekar and S. Patil, "Skin cancer detection and classification", 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI), 2017. Available: 10.1109/iceei.2017.8312419.

[5] I. A. OZKAN and M. KOKLU, "Skin Lesion Classification using Machine Learning Algorithms", Int J Intell Syst Appl Eng, vol. 5, no. 4, pp. 285–289, Dec. 2017. [6] D. Wen et al., "Characteristics of publicly available skin cancer image datasets: a systematic review", The Lancet Digital Health, vol. 4, no. 1, pp. e64-e74, 2022. Available: 10.1016/s2589-7500(21)00252-1.

[7] N. Hameed, A. Shabut, M. Ghosh and M. Hossain, "Multiclass multi-level classification algorithm for skin lesions classification using machine learning techniques", Expert Systems with Applications, vol. 141, p. 112961, 2020. Available: 10.1016/j.eswa.2019.112961.

[8] A. Murugan, S. Nair, A. Preethi and K. Kumar, "Diagnosis of skin cancer using machine learning techniques", Microprocessors and Microsystems, vol. 81, p. 103727, 2021. Available: 10.1016/j.micpro.2020.103727.

[9] C. Magalhaes, J. Tavares, J. Mendes and R. Vardasca, "Comparison of machine learning strategies for infrared thermography of skin cancer", Biomedical Signal Processing and Control, vol. 69, p. 102872, 2021. Available: 10.1016/j.bspc.2021.102872.

[10] M. Dildar et al., "Skin Cancer Detection: A Review Using Deep Learning Techniques", International Journal of Environmental Research and Public Health, vol. 18, no. 10, p. 5479, 2021. Available: 10.3390/ijerph18105479 [Accessed 17 August 2022].

[11] Y. Wang et al., "Deep learning enhances polarization speckle for in vivo skin cancer detection", Optics & Laser Technology, vol. 140, p. 107006, 2021. Available: 10.1016/j.optlastec.2021.107006 [Accessed 17 August 2022].

[12] R. Patil and S. Bellary, "Machine learning approach in melanoma cancer stage detection", Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 6, pp. 3285-3293, 2022. Available: 10.1016/j.jksuci.2020.09.002 [Accessed 17 August 2022].

[13] S. I. Ullah, A. Salam, W. Ullah, and M. Imad, "COVID-19 Lung Image Classification Based on Logistic Regression and Support Vector Machine," in European, Asian, Middle Eastern, North African Conference on Management & Information Systems, 2021: Springer, pp. 13-23.

[14] M. Imad, N. Khan, F. Ullah, M. A. Hassan, and A. Hussain, "COVID-19 classification based on

Chest X-Ray images using machine learning techniques," Journal of Computer Science and Technology Studies, vol. 2, no. 2, pp. 01-11, 2020.

[15] A. Hussain, M. Imad, A. Khan and B. Ullah, "Multi-class Classification for the Identification of COVID-19 in X-Ray Images Using Customized Efficient Neural Network", AI and IoT for Sustainable Development in Emerging Countries, pp. 473-486, 2022. Available: 10.1007/978-3-030-90618-4_23

[16] M. Imad, A. Hussain, M. Hassan, Z. Butt and N. Sahar, "IoT Based Machine Learning and Deep Learning Platform for COVID-19 Prevention and Control: A Systematic Review", AI and IoT for Sustainable Development in Emerging Countries, pp. 523-536, 2022. Available: 10.1007/978-3-030-90618-4_26

[17] M. Imad, F. Ullah, and M. A. Hassan, "Pakistani Currency Recognition to Assist Blind Person Based on Convolutional Neural Network," Journal of Computer Science and Technology Studies, vol. 2, no. 2, pp. 12-19, 2020.

[18] M. Imad, S. I. Ullah, A. Salam, W. U. Khan, F. Ullah, and M. A. Hassan, "Automatic Detection of Bullet in Human Body Based on X-Ray Images Using Machine Learning Techniques," International Journal of Computer Science and Information Security (IJCSIS), vol. 18, no. 6, 2020.

[19] M. Imad, S. I. Ullah, A. Salam, W. U. Khan, F. Ullah, and M. A. Hassan, "Automatic Detection of Bullet in Human Body Based on X-Ray Images Using Machine Learning Techniques," International Journal of Computer Science and Information Security (IJCSIS), vol. 18, no. 6, 2020.

[20] "Skin Cancer MNIST: HAM10000", Kaggle.com, 2022. [Online].Available: https://www.kaggle.com/datasets/kmader/skin-cancer-mnistham10000.979-8-3503-2750-2/23/$31.00 ©2023 IEEE