International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 09 | Sep 2024 www.irjet.net p-ISSN: 2395-0072

Towards Robust Skin Cancer Diagnosis: Deep Fusion of VGG16 and Mobile Net Features

Anirudh Kuchibhotla[1]

SRM University[1]

---***---

Abstract – *Skin cancer has become a dangerous condition in the modern world. It can be classified as either non-melanoma or melanoma (benign or malignant). Scars, dark spots, or changes in the skin appearance can all be indicators of skin cancer. Changes in irritability, size, form, or color could be indicators of skin cancer. This project concentrates on ensembling the extracted features from the vgg16 and MobileNet pre-trained models and fed to the ML models and a deep neural network. To enhance some current procedures and create modern approaches that would enable precise models that decrease the time gap between the diagnosis and treatment period in identifying skin cancer. With ensemble features from vgg16 and mobileNet pre-trained models, experimental findings show the effectiveness of the deep neural network and achieve an accuracy of 87%. Index*

***Key Words*:-** *Skin cancer, Malignant, Benign, Deep Neural Network,pre-trained models*

1.INTRODUCTION

Millions of people worldwide lose their lives to skin cancer every year, which is one of the most common types of cancer. Each year, this condition affects about three million people in the United States alone.[6] This is more than the total number of cases of breast, lung, and colon cancer combined together. The main cause of skin cancer is UV radiation exposure. Physicians typically perform a biopsy in order to make a diagnosis. Skin cancer, which includes both benign and malignant types, is one of the most important health issues facing nowadays. Planning an appropriate diagnosis and course of treatment requires the ability to distinguish between tumors that are benign and malignant. Early warning signs of skin cancer may be visible changes in the look of the skin, such as dark patches, scars, or texture changes. There are two main types of skin cancer: benign and malignant. Melanoma is one example of a malignant skin cancer that develops anywhere and is derived from melanocytes.

They have the ability to spread, aggressive growth patterns, and invasion of nearby tissues. Benign skin tumors, on the other hand, usually grow more slowly and do not behave in an aggressive manner. A correct diagnosis is essential to medicine since it guides treatment choices and determines prognosis. This highlights the importance of distinguishing between various sorts. These days, machine learning (ml) and deep learning (dl) help distinguish benign and malignant skin lesions. While technology might be beneficial to healthcare professionals, it shouldn't replace their training and expertise.

This study is focused on the detailed examination of skin lesions, including benign and malignant forms. The project's goal is to create reliable diagnostic techniques by closely analyzing the various traits that these lesions display. It also tackles the problem of improving the way that skin cancer is currently diagnosed by fusing cutting-edge technology with machine learning methods.In particular, it concentrates on combining features that are taken from trained models like VGG16 and MobileNet. These features are then fed into a deep neural network and a variety of machine learning algorithms by utilizing ensemble techniques. Notably, the obtained accuracy of 87% indicates promising developments in the field of skin cancer detection and provides assurance for more effective and efficient diagnostic methods in the future.

2.RELATED WORK

The findings of the data regarding skin cancer is entailed by exploring 10 research papers from 2013 to 2023. Aman Kamboj et al.[7] proposed a methodology for melanoma skin cancer detection that incorporated artifact removal, segmentation, feature extraction, and classification algorithms. It employed preprocessing techniques such as artifact removal through thresholding and segmentation using K-means in the HSV color space. Feature extraction involved asymmetry and color features, with classification performed using SVM and KNN algorithms. This study underscored the importance of feature combination and machine learning techniques in achieving accurate melanoma detection, with a focus on enhancing diagnostic capabilities through a fusion strategy for segmentation. Evgin Goceri[3] discussed an Automated skin cancer detection that relies on dl methods for classifying and detecting lesions , with pre-processing techniques like DullRazor for hair removal. Architectures such as GoogleNet, VGG16, ResNet50, and DenseNet are merged to enhance accuracy in distinguishing benign and malignant lesions. Challenges included high computational costs, driving the need for faster training methods for mo�bile device compatibility.Databases like ISIC, HAM10000, PH2, and ISBI2017 are utilized for training these models, emphasizing the importance of robust datasets in advancing skin cancer detection technology. Shagun Sharma et al.[11] explored machine learning and deep learning methods to classify skin cancer. Among SVM, KNN, NB, and NN models, NN emerged as the top performer. Highlighting the importance of early detection amidst increasing melanoma cases, this study evaluated models on a skin cancer dataset. NN demonstrated superior performance making it as the preferred choice for skin cancer classification. Mohamed Khalad Abu Mahmoud et al.[9] employed K-means clustering and spatially adjustable color median filtering , sequential feature selection, and SVM classification, the system achieves an 88.9 percent accuracy rate in diagnosing malignant melanoma. The method�ologies included feature extraction from the histograms and co-occurrence matrix, sequential feature selection using SVM, and K-means clustering for segmentation. The high sensitivity (87.5 percent) and specificity (100 percent) in skin tumor classification demonstrate the system's effectiveness in enhancing early detection and accuracy in skin cancer diagnosis. Muhammad Qasim Khan et al.[8] presented a system utilizing hybrid features integrating texture and color characteristics of skin lesions. The methodology involved pre-processing for image quality enhancement and ROI extraction for feature extraction. Various classification algorithms and image pro�cessing methods classified features into melanoma or nevus. Emphasizing GLCM, LBP, and Supervector for accurate skin cancer classification, the study underscored the importance of feature extraction techniques. Milon Hossain et al.[5] dis�cussed the application of ResNet models, including ResNet 18, 50, 101, and 152, for skin cancer diagnosis. It outlined the methodology of utilizing deep neural network models, particularly in PyTorch, for this purpose. The training process involved increasing epochs to enhance accuracy, with ResNet 152 demonstrating the best performance after 25 epochs.

The models leverage skip connections for improved results, and the output analysis involved class probability and predicted images to differentiate skin lesions. Mehwish Dildar et al.[1] carried out a comprehensive analysis of deep learning methods for detecting skin cancer, with an emphasis on classifying lesions through the use of radial basis function networks (RBFNs), k�nearest neighbors (KNNs), artificial neural networks (ANNs), and convolutional neural networks (CNNs). This study out�lined the research methodology, including search strings, cri�teria, and evaluation procedures. It highlighted the trends and challenges in skin cancer detection systems, emphasizing the need for further improvement in diagnostic techniques. Pronab Ghosh et al.[2] introduced The deep-learning model SkinNet�16 for distinguishing benign and malignant skin lesions. It incorporated optimization algorithms such as Adamax, image processing evaluations using MSE, PSNR, and SSIM values, and dl models for skin cancer detection. The study utilized automated medical image segmentation techniques, robust hair removal methods, and feature extraction. It highlighted the importance of segmentation techniques and deep learning methodologies in improving skin cancer diagnosis through advanced image analysis. The main limitations from the observed papers are: • Some studies may have limited access to large-scale datasets, resulting in small sample sizes for training and testing. This limitation can affect the robustness and reliability of the developed models, particularly in detecting rare or uncommon cases of skin cancer.

• While promising results are obtained in controlled experimental settings, the clinical validation of these models in real-world healthcare settings is limited. Clinical validation is essential to assess the practical utility, reliability, and safety of these systems when deployed in clinical practice.

• Deep learning models, while achieving impressive performance, frequently lack clarity, which makes it difficult to comprehend the underlying causes of their forecasts. Interpretability is crucial in medical applications to gain trust from healthcare professionals and ensure the transparency of the diagnostic process.

• Many advanced models, such as deep neural networks, re�quire significant computational resources for training and inference. This requirement may limit their accessibility, particularly in resource-constrained environments or regions with limited access to high-performance computing infrastructure.

3. METHODOLOGY

The progress of the project is intiated from the normalization of the images(i.e, Skin lesions of benign or malignant) and ensembled features from vgg16 and mobileNet pre-trained models are fed to the deep neural Network. VGG16, with its deep and rich architecture, excels at capturing intricate patterns and details in images, making it highly effective for tasks requiring fine-grained image classi�fication. On the other hand, MobileNet's lightweight design and efficiency make it well-suited for applications where computational resources are limited, without sacrificing too much on accuracy. By leveraging the feature representations learned by each model, ensembling aims to capture a deeper comprehension of the input data that results in increased robustness and accuracy in predictions.

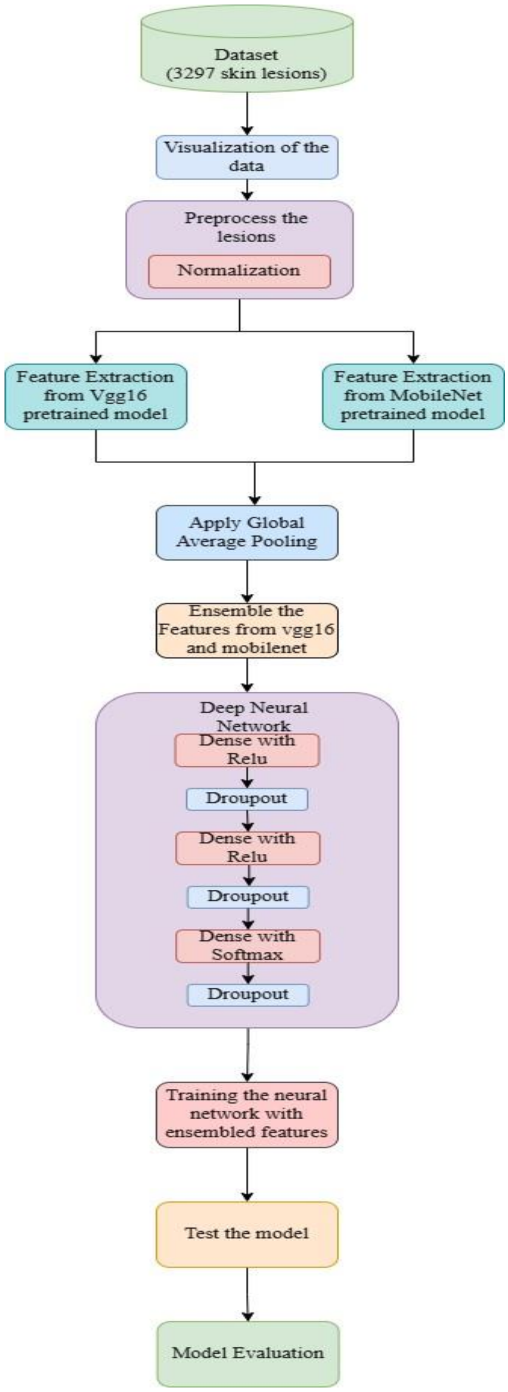

Fig-1:Methodology

3.1 Dataset Skin

Lesions of benign and malignant are taken from the source Kaggle,which is a large collection of Datasets and this dataset contains 3297 images of both Benign and Malignant and the data is already divided into train and validation 2637 and 660 images respectively.

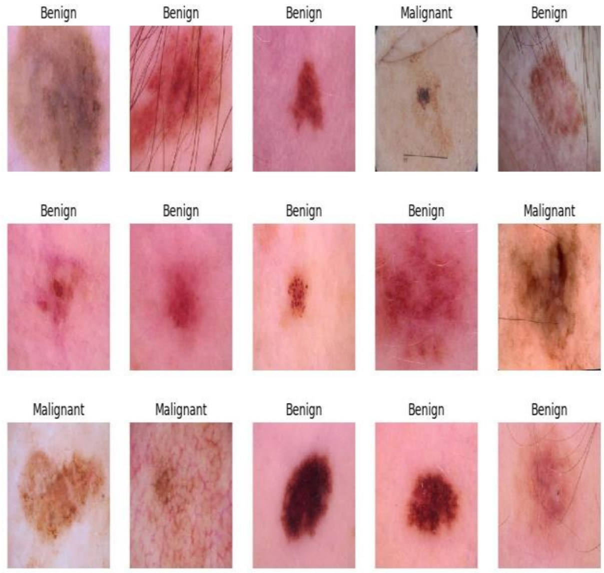

Fig-2:A Few Sample Pictures from the Dataset

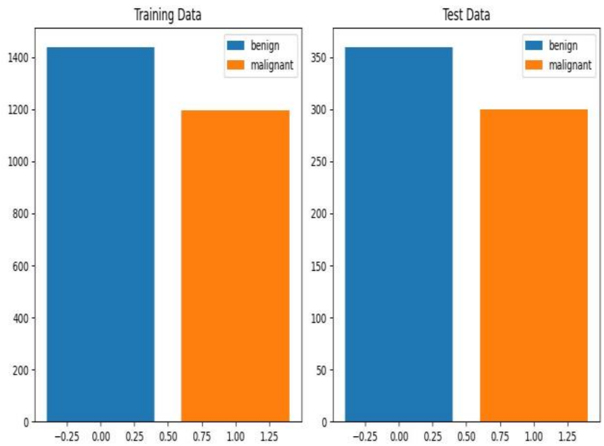

Fig-3:Distribution of Train and Test Images

3.2 Image Preprocessing Normalization:

Normalization is crucial for ensuring fair and accurate model training and evaluation. Skin cancer datasets often exhibit class imbalances between benign and malignant samples, with one class significantly outnumbering the other. Normalization techniques address this issue by adjusting the distribution of samples across classes, thereby preventing the model from being biased towards the majority class and ensuring equitable representation of both benign and malignant lesions during training. Normalization in skin cancer classification refers to balancing the proportion of malignant and benign samples in the training dataset.

3.3 Feature Extraction:

1.Vgg16: The deep convolutional neural network architecture known as VGG16 emerged to fame due to its superior performance in classification of images applications. VGG16, created by the University of Oxford's Visual Geometry Group, is distinguished by its uniform construction and simplicity. It has sixteen layers total, the majority of which are max-pooling and convolutional layers. Due to its pre-training on the extensive ImageNet dataset, VGG16 is capable of learning rich feature representations from a diverse range of images. [4] Due to its straightforward design and impressive performance, VGG16 has become a popular choice for transfer learning and feature extraction in computer vision tasks.

2.MobileNet: A lightweight convolutional neural network architecture called MobileNet was created especially for embedded and mobile devices with constrained computational power. Using depthwise separable convolutions, Google's MobileNet model strikes a compromise between accuracy and size. The net�work's capacity to extract key features from input images is preserved by these convolutions, which also dramatically lower the computational cost. MobileNet is especially well-suited for workloads requiring real-time inference on mobile devices since it has been pre-trained on huge datasets like ImageNet. Particularly in contexts with limited resources, MobileNet's small size and effectiveness make it a strong option for a variety of applications, such as semantic segmentation, object detection, and image classification. This research uses VGG16 and MobileNet as a potent tool to extract significant features from the dataset images, as opposed to training them from scratch. The process involves passing the normalized images through the VGG16 model and MobileNet, which analyzes them and identifies important patterns and structures. The output of this analysis is a set of numerical features that encapsulate key aspects of each image. By doing this for both our training and testing datasets, new sets of features are obtained.

3.4 Global Average Pooling:

Global Average Pooling is applied to the feature maps extracted from two different pre-trained models: VGG and MobileNet. This process condenses the feature representations while preserving crucial spatial information. It involves computing the average value of each feature map across its spatial dimensions, typically height and width. This reduces the spatial dimensions to 1x1 and yields a single value for each feature map.

$$\bar{f}_c = \frac{1}{A \times B} \sum_{i=1}^{A} \sum_{j=1}^{B} f_{ijc}$$

where A denotes Height, B denotes Width, fijc denotes the activation value at spatial location (i, j) for channel c. This calculation is performed for each channel independently, resulting in a 1x1xC tensor.

3.5 Ensemble the Extracted Features of Vgg16 and MobileNet:

Ensembling the extracted features of VGG16 and MobileNet pre-trained models involves combining the feature representations obtained from these two distinct convolutional neural network architectures. Each pre-trained model extracts features from input images, capturing different aspects of the visual information due to their diverse architectures and training procedures.By ensembling these features, we aim to leverage the complementary strengths of both models, enhancing the overall performance of the system.By reducing individual model biases and errors, it can produce predictions that are more solid and trustworthy.

3.6 Architecture of Deep Neural Network:

The defined Sequential neural network model comprises multiple layers aimed at effectively learning from the ensem�bled features of vgg16 and MobileNet and making accurate predictions. The first layer is dense, with 128 neurons triggered by the Rectified Linear Unit (ReLU) function. This is followed by a dropout layer, which deactivates neurons at random during training to prevent overfitting, and has a dropout rate of 0.5. The next set of layers exhibits a like structure, comprising an additional dense layer with 64 neurons and a dropout layer, succeeded by a dense layer with 32 neurons and a matching dropout layer. To facilitate binary class classification tasks, the final output layer generates probability scores for each class using the softmax activation function. This architecture aims to strike a balance between model complexity and generalization capability, ensuring robust performance on unseen data while effectively leveraging the features extracted from the ensembled VGG16 and MobileNet pre-trained models.

3.6 Model Training:

Concatenated features from the pre-trained VGG16 and MobileNet models are used to train the model, along with the matching training labels. The loss function, which calculates the difference between the predicted and real labels, is minimized by the model by adjusting its internal parameters (weights) during the training process. The model learns from the training data during a series of iterations, or epochs, in the training process. in batches. In each epoch, the model receives batches of training data, computes the loss, and updates the weights using optimization techniques such as Adam optimization. An early halting callback is used to keep an eye on the model's performance and prevent overfitting. This callback restores the best weights from training and stops training if the validation accuracy does not improve after a predetermined number of epochs (patience). The overall goal of the training process is to maximize the model's parameters such that, using the features that were retrieved from the pre-trained models, the skin lesion images may be reliably classified as benign or malignant.

3.7 Model Testing:

A different dataset, referred to as the test set, is used in the testing procedure. It contains visuals that the model did not see during training. During testing, proposed model takes the concatenated features extracted from the VGG16 and MobileNet pre-trained models as input and predicts the corresponding labels for the test samples. These predicted labels are then compared with the ground truth labels to compute performance metrics such as accuracy, precision, recall, and F1 score.

3.8 Model Evaluation:

An overall performance of the model is provided by the test accuracy, which is calculated as the fraction of correctly identified samples out of the overall number of test samples. Furthermore, other measures like precision and recall provide information about how well the model can distinguish between benign and malignant lesions and prevent misclassifications. This research evaluates the model's generalization capabilities with data that hasn't been seen before and determine its effectiveness in accurately identifying skin cancer lesions. The proposed achieved an accuracy of 87%.

4. EXPERIMENTAL RESULTS

About Dataset: The dataset named skin cancer benign vs malignant is souced from the kaggle datasets, which consists of train and test images of total 3297 images. It has 2 classes named benign and malignant in separate folders for both train and test. Later,the dataset is trained on various pretrained models like Vgg16,Vgg19,ResNet50,MobileNet and GoogleNet. Evaluation of pre-trained models are recorded in table I.

TABLE I
EVALUATION METRICS FOR PRE-TRAINED MODELS

Model	Classes	F1 Score	Precision	Recall	Accuracy
VGG-16	Benign Malignant	0.87 0.85	0.88 0.84	0.86 0.86	86.2
VGG-19	Benign Malignant	0.85 0.83	0.88 0.80	0.82 0.86	84.09
GoogLeNet	Benign Malignant	0.77 0.78	0.87 0.70	0.69 0.87	77.42
MobileNet	Benign Malignant	0.86 0.85	0.90 0.81	0.82 0.90	85.45
ResNet	Benign Malignant	0.84 0.80	0.82 0.78	0.85 0.80	82

From this,Vgg16 and MobileNet models outperformed well with 86.2% and 85.45% respectively when compared to pro�posed model in [10]. So, the features from these two models are ensembled and fed to the ML models in table V and a neural network of 3 dense layers and 3 droupout layers. From Table II, Deep neural network achieved an accuracy of 87.4% among all the models.

TABLE II
EVALUATION METRICS FOR DIFFERENT MODELS

Model	Classes	F1 Score	Precision	Recall	Accuracy
Logistic Regression	Benign Malignant	0.86 0.83	0.87 0.82	0.84 0.84	84.09
Random Forest	Benign Malignant	0.85 0.83	0.87 0.81	0.83 0.85	84.09
Decision Tree	Benign Malignant	0.78 0.75	0.80 0.73	0.77 0.76	76.51
SVM	Benign Malignant	0.84 0.81	0.86 0.79	0.82 0.84	82.72
Deep Neural Network	Benign Malignant	0.88 0.86	0.88 0.86	0.89 0.86	87.4

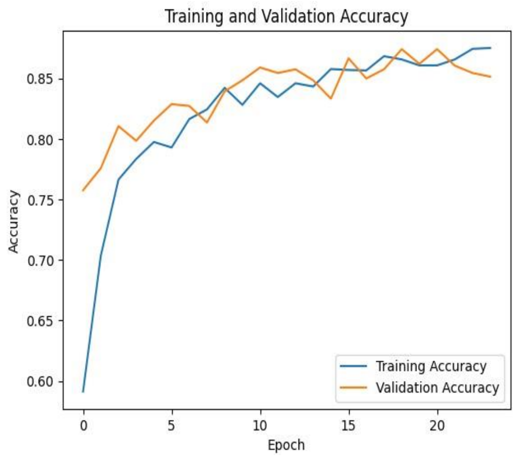

Fig-4:Proposed Model's accuracy in training and validation

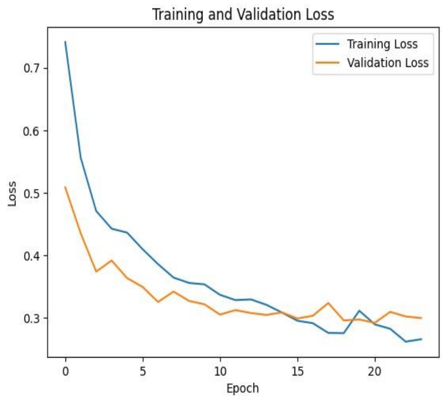

Fig-5:Training and Validation Loss of Proposed Model

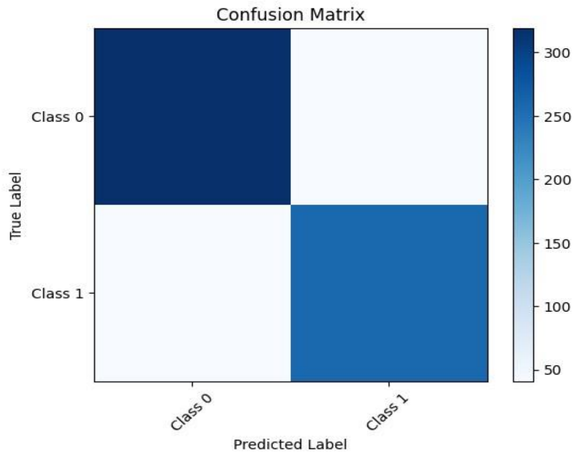

Fig-6:Confusion Matrix of Proposed Model

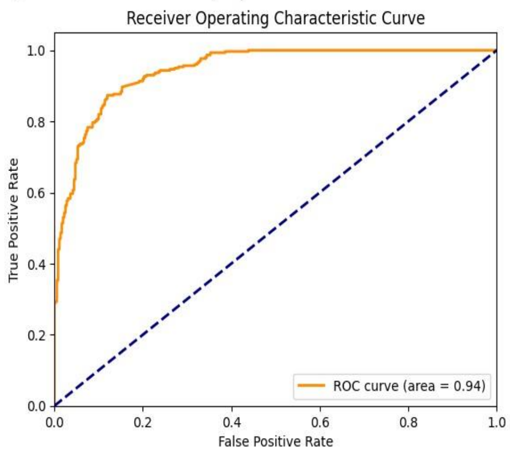

Fig-7:ROC Curve of Proposed Model

5. CONCLUSION

This project evaluated the suggested system's performance rigorously using metrics like accuracy, precision, recall, and F1 score. The experimental results shows how well the deep neural network performed effectively when combined with ensemble features from VGG16 and MobileNet pre-trained models. Achieving an accuracy of 87% demonstrates the promising potential of the approach in accurately detecting skin cancer lesions. By integrating cutting-edge technologies with medical diagnostics, this project strive to contribute to the development of more efficient and reliable methods for combating skin cancer, ultimately improving patient outcomes and saving lives.

6. REFERENCES

[1] *Jayaram Nori*[*1] [*1]Broadcom Inc, USA. DOI :- https://www.doi.org/10.56726/IRJMETS53503

[2] *P. Prasad and A. Rao, "A survey on various challenges and issues in implementing AI for enterprise monitoring," Journal of Network and Computer Applications, vol. 116, pp. 42-55, 2018, doi: 10.1016/j.jnca.2018.05.005.*

[3] *Y. Dang, Q. Lin, and P. Huang, "AIOps: real-world challenges and research innovations," in 2019 IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings (ICSE-Companion), 2019, pp. 4-5, doi: 10.1109/ICSE-Companion.2019.00023.*

[4] *D. Xu et al., "Unsupervised anomaly detection via variational auto-encoder for seasonal KPIs in web applications," in Proceedings of the 2018 World Wide Web Conference, 2018, pp. 187-196, doi: 10.1145/3178876.3185996.*

[5] *V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," ACM Computing Surveys (CSUR), vol. 41, no. 3, pp. 1-58, 2009, doi: 10.1145/1541880.1541882.*

[6] *M. Chen et al., "Big data: A survey," Mobile Networks and Applications, vol. 19, no. 2, pp. 171-209, 2014, doi: 10.1007/s11036-013-0489-0.*

[7] *Y. Li et al., "Deep learning for anomaly detection in cloud native systems," in 2020 IEEE International Conference on Cloud Engineering (IC2E), 2020, pp. 106-116, doi: 10.1109/IC2E48712.2020.00022.*

[8] *F. Salfner, M. Lenk, and M. Malek, "A survey of online failure prediction methods," ACM Computing Surveys (CSUR), vol. 42, no. 3, pp. 1-42, 2010, doi: 10.1145/1670679.1670680.*

[9] *F. Jiang et al., "Artificial intelligence in healthcare: Past, present and future," Stroke and Vascular Neurology, vol. 5, no. 2, 2020, doi: 10.1136/svn-2020-000443.*

[10] *F. Salfner, M. Lenk, and M. Malek, "A survey of online failure prediction methods," ACM Computing Surveys (CSUR), vol. 42, no. 3, pp. 1-42, 2010, doi: 10.1145/1670679.1670680.*

[11] *X. Liu et al., "PANDA: Facilitating usable AI development," arXiv preprint arXiv:2003.04070, 2020.*

[12] *G. A. Susto, A. Beghi, and C. De Luca, "A predictive maintenance system for epitaxy processes based on filtering and prediction techniques," IEEE Transactions on Semiconductor Manufacturing, vol. 25, no. 4, pp. 638-649, 2012, doi:10.1109/TSM.2012.2209131.*

[13] *X. Liu et al., "PANDA: Facilitating usable AI development," arXiv preprint arXiv:2003.04070, 2020.*

[14] *E. Cortez et al., "Resource central: Understanding and predicting workloads for improved resource management in large cloud platforms," in Proceedings of the 26th Symposium on Operating Systems Principles (SOSP), 2017, pp. 153-167, doi: 10.1145/3132747.3132772.*

[15] *Z. Yin et al., "An empirical study on configuration errors in commercial and open source systems," in Proceedings of the 26th Symposium on Operating Systems Principles (SOSP), 2017, pp. 159-176, doi: 10.1145/3132747.3132773.*

[16] *D. Wang et al., "Failure prediction using machine learning in a virtualised HPC system and application," Cluster Computing, vol. 20, no. 1, pp. 103-115, 2017, doi: 10.1007/s10586-016-0668-4.*

[17] *J. Gao, "Machine learning applications for data center optimization," Google White Paper, 2014. [Online]. Available: https://research.google/pubs/pub42542/*

[18] *R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in 2010 IEEE Symposium on Security and Privacy, 2010, pp. 305-316, doi: 10.1109/SP.2010.25.*

[19] *R. Boutaba et al., "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities," Journal of Internet Services and Applications, vol. 9, no. 1, pp. 1-99, 2018, doi: 10.1186/s13174-018-0087-*

[20] *A. Mestres et al., "Knowledge-defined networking," ACM SIGCOMM Computer Communication Review, vol. 47, no. 3, pp. 2-10, 2017, doi: 10.1145/3138808.3138810.*