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

AGCRNL: Automatic Glaucoma Classification Using Residual Network and LSTM

Jaswanth N [1]

[1]SRM University

Abstract - *In this work, the diagnosis of glaucoma in the early stage using deep learning models has been presented. Glaucoma is a group of diseases which is irreversible and can cause blindness. Manual diagnosis depends on human skills and it is difficult to diagnose at an early stage. Glaucoma classification using deep learning is a challenging task. The concept of feature extraction using deep learning models has been utilized in the current study. An ensemble residual network along with squeeze and excitation block and with BiLSTM (Binary Long Short-Term Memory) has been implemented for sequential feature extraction for faster and more accurate classification of the disease. A residual network is used for global feature extraction. The accuracy rate of the validation set was 92.3% for our model. The findings suggest that a cost-effective screening tool for early and cost-effective identification of glaucoma could be developed utilizing deep learning algorithms.*

Key Words: *Bi-LSTM, Glaucoma Classification, ResNet50, SE block.*

1.INTRODUCTION

Glaucoma is the second progressive worldwide leading cause of eye diseases. The early-stage detection of glaucoma is a challenging task. Currently, around 79.6 million individuals are affected by glaucoma [1] and it has been estimated that by 2040 the number of affected individuals will be around 111.8 million [2]. The eye is one of the most complex, sensitive and delicate organs which is responsible for passing visual information to the brain cells in the human body.

The root cause of glaucoma is the elevated pressure in the eye which can be associated with Intra-Ocular Pressure (IOP) [3]. Improper drainage of the liquid aqueous humor leads to an increase in pressure in the eye subsequently affecting the retina and optic nerve and human vision. In general, the IOP value should be less than 21mmHg to fight against glaucoma. Manual diagnosis is based on structural changes of the retinal nerve fibre layer and optic nerve head [4]. Primarily, glaucoma can broadly be classified into two categories: 1) Open-angle glaucoma and 2) Angleclosure glaucoma. The effect of Open-angle glaucoma disease is not realized until it starts to impair human vision [2]. This type of glaucoma disease is painless and is caused when the drainage canal is not able to drain out the excess aqueous humor. People with high blood pressure and diabetes are more susceptible towards this type of glaucoma. In the Angle-closure glaucoma disease, the iris tends to block the drainage system by being extremely close to it. It requires immediate attention and can be severely painful.

Manual analysis of glaucoma can be time-consuming and cumbersome. The accuracy of parameter measurement for detection varies according to the doctors. In recent years, various automated techniques and diversity indices have been proposed for the automatic detection of the disease using retinal fundus images [5]. The Residual Network 50 (ResNet50) model has proven effective for the extraction of features and classifying the fundus image [6]. The accuracy of the model was found to be 93%. The residual network has been used in numerous research for feature extraction including detection of colorectal cancer and insect pest recognition [7, 8]. Taking motivation from the research of Ovreiu et. al. (2020) [9], Residual with squeeze and excitation block has been combined in this research to obtain channel-wise recalibrated features. It reduces and discards features based on channel re-calibration and helps to train the model in an efficient way.

In this article, we propose a novel architecture for fully automated binary classification of glaucoma where the Fundus database has been utilized for training and testing purposes. Different databases (e.g., ORIGA, ACRIMA and Fundus) can also be used for glaucoma classification. In the proposed framework, a residual network has been combined with a squeeze and excitation block followed by a BiLSTM block [10]. In addition, different available methods for glaucoma classification have been compared with the proposed model.

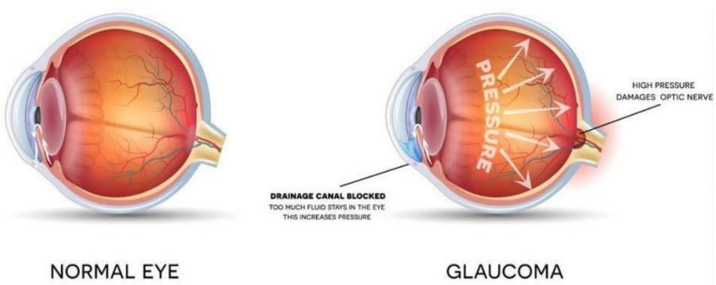

Fig -1: *Diagram of normal eye and eye with glaucoma [11]*

MOTIVATION

Glaucoma disease is the second leading cause of blindness globally [12]. The critical work of identifying and categorizing glaucoma is challenging which motivated us to initiate this project. Early detection and treatment can stop irreversible vision loss and enhance the quality of life for those who are affected. Advances in deep learning and artificial intelligence approaches for glaucoma categorization are a major influence on the lives of millions of people at risk of glaucoma [5]. These techniques also provide intriguing potential for early identification and effective monitoring of the condition while preserving vision, enhancing patient lives, and advancement in ophthalmology.

OBJECTIVE

The main objective of this study can be summarized as follows:

- The proposed model uses an ensemble method that deals with non-homogeneous and non-linear data elements while focusing on producing more accurate results as compared to models with a single level of feature extraction.
- Global feature extraction using the residual network is combined with squeeze and excitation block to perform dynamic channel-wise feature re-calibration. Sequential feature extraction using BiLSTM helps to extract every temporal feature of the image with accurate details.

LITERATURE REVIEW

Previous researchers have defined many embedded models for feature extraction and classification. In 2007, Boland and Quigley [1] classified risk factors into six categories that are most likely to lead to the development of open-angle glaucoma in their study. In 2009, a major addition to the field of ophthalmology was made through the study paper by Nayak et. al. [2] which describes an automated process for feature extraction and SVM for classification. According to the authors, the optic nerve head (ONH) and retinal nerve fiber layer (RNFL) features that are known to be associated with glaucoma can be extracted using digital fundus images. Ramani et. al. (2012) [3] described a novel method for classifying normal, glaucoma, and diabetic retinopathy cases using retinal image analysis and data mining methods. The authors strongly emphasize using computational methods to automatically forecast these eye diseases which can be done by utilizing cutting-edge image analysis and data mining techniques. The Grid Color Moment method technique was proposed by Ghosh et. al. (2015) [4] for obtaining color information from retinal pictures used in glaucoma classification. The image was divided into a grid of smaller sections and color moment characteristics were calculated within each grid cell. The characteristics defining the statistical distribution of color values were subsequently fed into a neural network classifier. This approach demonstrated encouraging results in reliably distinguishing normal and glaucomatous retinal pictures potentially aiding in the early detection and diagnosis of glaucoma. Chakravarty and Sivaswamy (2016) [5] introduced a glaucoma classification method that integrated segmentation- and imagebased features for binary classification of glaucomatous and normal retinal pictures. The researchers used both segmentation-based and image-based approaches to extract characteristics from retinal pictures. In 2016, Dey and Bandyopadhyay [6] proposed a support vector machine (SVM) classification method for automated glaucoma detection. Principal component analysis (PCA) was used for feature extraction and SVM was used for classification. PCA was utilized to decrease the extracted features and the smaller feature set was then used as input for SVM classification. The proposed method detected glaucoma with an amazing 96% accuracy rate. Hu et. al. (2018) [8] introduced the SE block in their proposal for squeeze-and-excitation (SE) networks which models channel dependencies and automatically calibrates channelwise feature responses. By adaptively weighting feature maps based on their channelwise relevance, the SE block is a small and effective method that collects channelwise information. Ferreira and Vinicius (2018) [13] proposed a model using a convolution neural network and SVM classifier by using textual descriptors. The CNN extracts the features and the SVM stage was used for classification. A hybrid feature space for glaucoma classification that integrates texture data with transfer learning utilizing deep learning algorithms was proposed by Claro et. al. (2019) [7].

The work made use of transfer learning, which uses deep learning models that have already been trained to extract important features from big datasets, together with texture-based features that were recovered from retinal pictures. The classification of glaucoma using the hybrid feature space created by fusing texture information and transfer learning produced good results. The research paper entitled “Early Detection of Glaucoma Using Residual Networks” by Ovreiu et. al. (2020) [9] used residual networks to extract features and classify images into normal or glaucoma which had an accuracy of around 93%. Bisneto et. al. (2020) [12] proposed a model which uses SVM and generative adversarial networks. The accuracy of the model was found to be 87.5%. Janani and Rajamohana (2021) [11] gave a review of numerous methods for the segmentation of the optic disc and optic cup for detecting glaucoma at an early stage. Oza et. al. (2021) [14] proposed a model for classification using a convolution neural network. The accuracy of the model was found to be 95.44%. In 2022, ODGNet: a deep-learning model for automatic optic disc localization and glaucoma classification using fundus images was proposed by Latif et. al. [10]. The research introduces the ODGNet deep learning model, which uses fundus pictures to automate the localization of the optic disc and the categorization of glaucoma. The suggested model uses deep learning to identify the optic disc in fundus pictures and categories of glaucoma cases [10]. In recent days, a number of attempts has made for automatic classification of glaucoma [24, 25, 26].

MATERIALS AND METHODS

EXISTING MODEL

In the existing model [9], the Residual network used for classification achieved a classification accuracy of 93.28% for the validation data where a total of 4 convolution blocks, each consisting of 3 convolution layers are present. The Fundus database is used for training and validation purposes in this model.

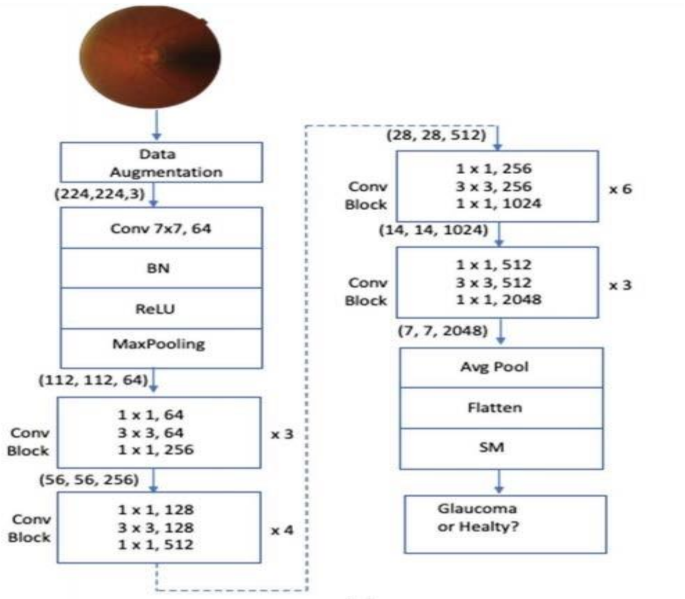

Figure 2: *Existing residual network model [9]*

PROPOSED MODEL

Four crucial components make up the glaucoma detection strategy in use.

Step 1: Data Augmentation and Pre-processing

Step 2: Global feature extraction: Residual network in combination with squeeze and excitation block (SE).

Step 3: Sequential feature extraction: BiLSTM

Step 4: Classification

The Image data generator is used to synthesize data using data augmentation and image quality is improved by re-sizing. The target image size is 224x224x3. For global feature extraction, the images are passed to the residual block and SE block. It extracts 2048 features for each image. The 2D image is flattened to serve as input to the long short-term memory for sequential feature extraction. Global average pooling is applied to combine the effect of 2048 features in a single feature for classification. A fully connected layer is used for the classification of images. Figure 3 shows the block diagram of the proposed model.

Step 1:

Data-Augmentation

Data Augmentation is a technique of generating synthetic data by rotating, scaling and shuffling the original data images. This technique is used to fulfil the requirements, the varied nature of the training data and the volume of data. In addition to these, the class imbalance issue in classification jobs can also be solved using augmented data. Models can be trained accurately and efficiently in deep learning by training with augmented data.

Since deep learning models require a large amount of data for training and validation, ImageDataGenerator has been employed to generate synthetic data for training and validating the proposed model where the target image size is maintained at 224x224x3. In the proposed methodology from the Fundus dataset, 554 images have been utilized for training and 130 images for validation.

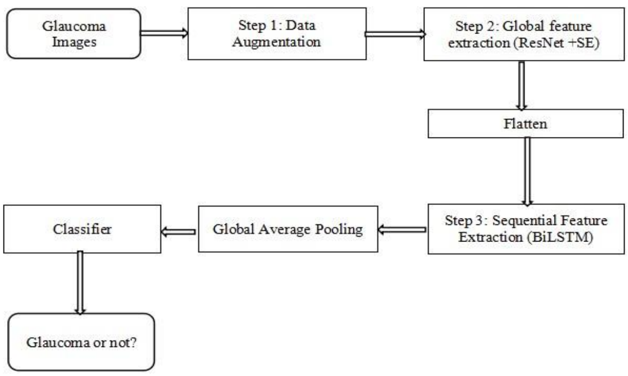

Figure 3: *Proposed model*

Step 2:

Global feature extraction

Global feature is a term used for the entire image and attributes associated with it. Feature extraction is a part of dimensionality reduction to reduce the variables and computing resource usage. A residual network is used for feature extraction in deep learning. Residual block provides a skip connection to minimize the vanishing gradient problem which can be considered as a striking advantage. ResNet50 architecture has an extra convolution layer to reduce the dependencies, number of parameters and matrix multiplication using bottleneck design for building blocks. ResNet 50 architecture has also been used for malicious software classification.

Squeeze and excitation blocks have two stages of action. SE block in the first 'Squeeze' step applies global average pooling to aggregate the feature map across the spatial dimension. An image having H x W x C is aggregated to 1 x 1 x C. In the excitation stage, the channels are reduced by a factor R = 16 (reduction ratio) according to per-channel weights. The channels are re-calibrated and using activation functions the channels are mapped to the output image to maintain the consistency in the number of channels [8]. SE block helps to discard the features which are not important for image classification and also it calculates the weight of each channel towards the classification during the training process.

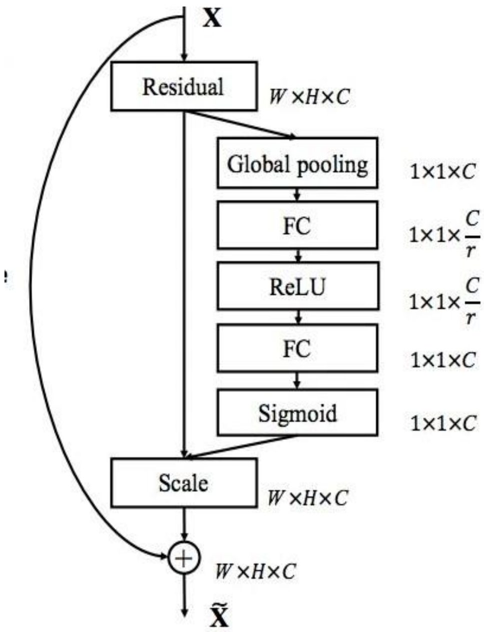

Figure 4: Residual network and SE Block

SE block combined with the residual block provides the same accuracy with much lower computational cost. The input for our model is 224x224x3 and the output of the global feature extraction block is in the form of 55x55x2048. 2048 features were extracted in step 2 for each image. Figure 4 and Figure 5 represent the block diagram of the residual network in combination with the SE block.

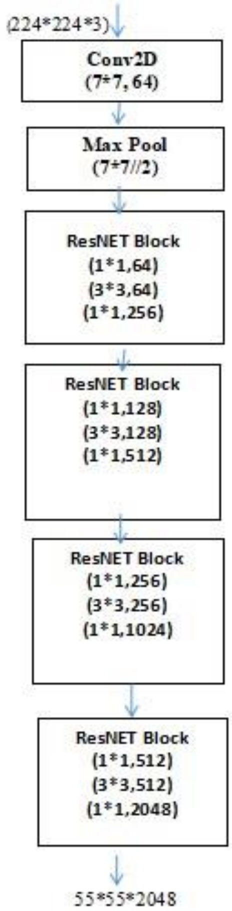

Figure 5: Model of global feature extraction

Step 3:

Sequential feature extraction

An additional BiLSTM model has been ensembled that takes the global features as input and generates the sequential features which are more accurate from both directions [10]. BiLSTM learns the features while extracting sequential features from the global ones. Each BiLSTM block or cell is responsible for the extraction of a single global feature which also focuses on local patterns and features of each global feature extracted. A total of 2048 BiLSTM cells are being used to cover 2048 global features.

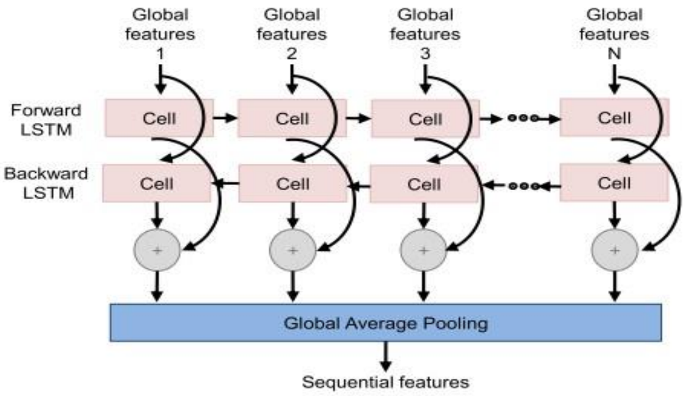

Figure 6: Sequential feature extraction [10]

LSTM has been utilized to store long sequences of data for a particular amount of time or until the forget gate is activated thereby reducing the long-term dependencies as it can remember the specific sequence and can be used for classification with reduced time. During the training process, it also learns the features and allows us to classify data with the least possible time.

Step 4: Classification

A dense layer or fully connected layer is placed after the global average pooling layer. It takes the aggregate feature map of 2048 features generated by the global average pooling and produces an output for binary classification where normal images are denoted as 1 and a glaucoma-affected image is denoted as 0.

RESULTS AND DISCUSSION

In the proposed model, global features were extracted using a residual block which was then followed by sequential feature extraction. Classification is done by a single dense layer. The results shown in Figures 8 and 9 are the result of 50 epochs.

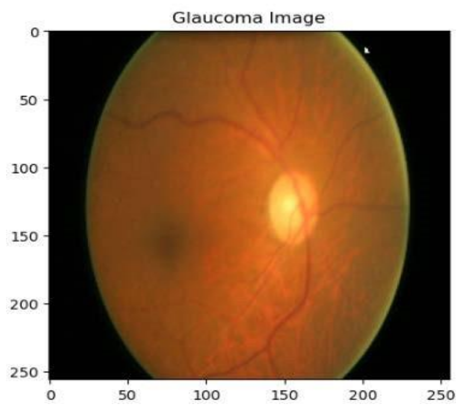

Figure 7: Original Image

Global Features

Residual block in combination with SE block takes the input image of size 224x224x3 and extracts 2048 features for each in the output format 55x55x2048. After flattening the global features were fed to the BiLSTM block for sequential feature extraction. Figure 8 represents 8 features from the 2048 extracted features.

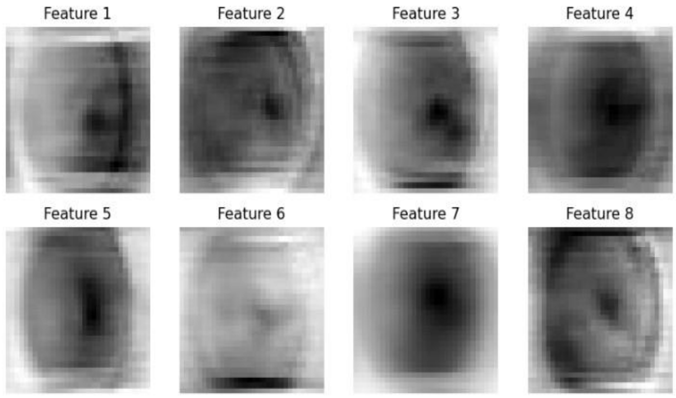

Figure 8: Features extracted Residual block

Sequential features:

BiLSTM block takes the global features generated by the residual block and focuses locally on each feature to generate a sequential feature or pattern. Since BiLSTM can work on linear data, flatten is applied to change the dimension of the features generated by the residual block. Figure 9 shows the features after passing each global feature from the BiLSTM block.

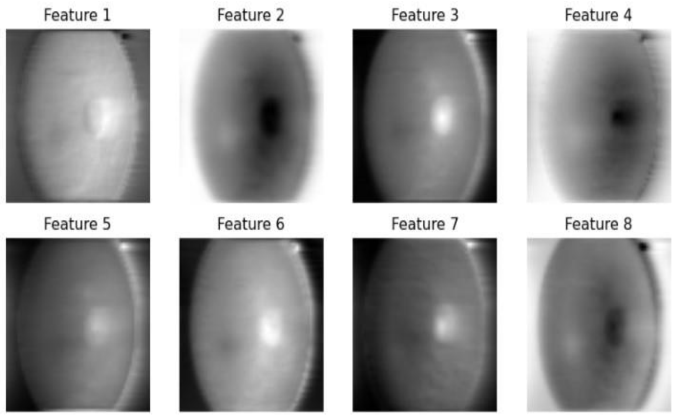

Figure 9: Features after BiLSTM

Quantitative Result and qualitative result

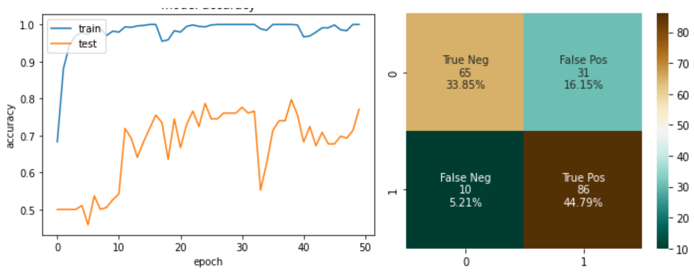

Figure 10: Training and validation accuracy curve and confusion matrix

Figure 10 shows the training and validation accuracy curves and the confusion matrices for the performance of the proposed model on the glaucoma datasets. These figures provide insight into how the accuracy of the model changes as the training epoch progresses and how the model performs on different classes of images. The accuracy curves shows that the performance of the proposed model improves significantly with increasing training epochs, while the confusion matrices provides a visual representation of the prediction made by the model as compared to the ground truth labels.

Table 2: Comparative analysis of proposed techniques in terms of various metrics with different datasets

Dataset	Architecture	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)
ACRIMA	**Proposed Method**	**99**	**96**	**96**	**95**
FUNDUS	**Proposed Method**	**98**	**95**	**92**	**94**
ORIGA	**Proposed Method**	**86**	**84**	**98**	**97**

Table 2 summarizes the performance of the proposed model for glaucoma classification using three different datasets, i.e., ACRIMA, FUNDUS and ORIGA.

The proposed model is based on an advanced ensemble-based CNN architecture called BiLSTM along with ResNet50 and SE block, which has shown promising results in accurately detecting glaucoma from retinal fundus images. The results of the study demonstrate that the proposed method achieves higher accuracy as compared to the existing methods for all three datasets, with an average accuracy of 94%. The precision of the proposed model is also high and ranges from 84% to 96%, indicating that the model has a low false positive rate. The recall rate, which measures the proportion of positive cases that are correctly identified by the model, ranges from 92% to 98%, indicating that the model also has a low false negative rate.

Table 3: Proposed model is evaluated comparatively with various existing methods

Architecture	Precision (%)
CNN	94
ResNet 50	93.28
Transfer Learning	94
CNN with texture descriptor	92
Resnet50 + SEblock+ BiLSTM	**98**

Table 3 shows the comparative analysis for the performance of the proposed model to other architectures used in different literatures. The architectures include CNN, ResNet50, Transfer Learning, and the proposed ResNet50 with SE block and BiLSTM model. The results demonstrates that the proposed model achieved the highest precision at 98% thereby outperforming all other architectures in the comparison. These findings suggest that the proposed model has the potential to be a valuable tool for accurately detecting glaucoma and preventing vision loss.

CONCLUSION

Glaucoma is a serious eye condition that can cause vision loss and even blindness if not detected early. To address this issue, a recent study proposed the use of ensemblebased CNN architectures, specifically BiLSTM with ResNet50 and SE block, as glaucoma classifiers. The study used only publicly available image datasets and data augmentation techniques to train and evaluate the performance of the proposed model. The results showed that the proposed architecture achieved the best performance, with an average accuracy of 0.98, an average specificity of 0.94, and an average recall of 0.92. The study also found that fine-tuning the proposed architecture led to these impressive results, which were achieved after analyzing 2594 images and implementing a sampling mechanism. Overall, these findings demonstrate the potential of using advanced CNN architectures to improve glaucoma diagnosis and ultimately prevent vision loss and blindness.

REFERENCES

[1] Jayaram Nori*1 *1Broadcom 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.