iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: VII Month of publication: July 2026

DOI: https://doi.org/10.22214/ijraset.2026.84316

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Machine Learning Approach for Calculating Healthcare Protection Premiums

D. Rishika[1], Tulasi Miryala[2]
[1]*Student, Master of Computer Applications (MCA), AQJ Centre for P.G. Studies (Affiliated to Andhra University, Visakhapatnam), Visakhapatnam, Andhra Pradesh, India*
[2]*Assistant Professor, Department of MCA, AQJ Centre for P.G. Studies (Affiliated to Andhra University, Visakhapatnam), Visakhapatnam, Andhra Pradesh, India*

Abstract: Healthcare insurance plays a vital role in reducing the financial burden associated with medical treatments and improving access to quality healthcare services. Determining an appropriate healthcare protection premium is a challenging task because it depends on multiple demographic, medical, and lifestyle factors that influence an individual's health risk. Conventional premium calculation methods generally rely on actuarial models and statistical techniques, which often fail to capture the complex nonlinear relationships among these variables. Consequently, insurance providers may encounter inaccurate premium estimation, affecting both customer satisfaction and financial stability.

This research proposes a machine learning-based framework for predicting healthcare protection premiums using Multiple Linear Regression (MLR), Random Forest Regressor (RFR), and Support Vector Regression (SVR). The proposed approach utilizes historical healthcare insurance data to identify patterns and relationships among customer attributes such as age, gender, Body Mass Index (BMI), smoking habits, medical history, policy duration, and family size. Data preprocessing techniques, including handling missing values, feature encoding, and normalization, are applied before model training. The models are evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2 Score). Comparative analysis indicates that the Random Forest Regressor provides the most accurate and reliable premium predictions. The proposed system improves premium estimation accuracy, reduces manual effort, supports personalized pricing, and enhances decision-making for healthcare insurance providers.

Keywords: Healthcare Insurance, Premium Prediction, Machine Learning, Multiple Linear Regression, Random Forest Regressor, Support Vector Regression, Predictive Analytics, Healthcare Protection.

I. INTRODUCTION

Healthcare insurance has become an essential component of modern healthcare systems because it provides financial protection against unexpected medical expenses. Rising treatment costs, increasing healthcare awareness, and the growing prevalence of chronic diseases have significantly increased the demand for reliable health insurance policies. Insurance providers must accurately estimate healthcare premiums to maintain financial sustainability while offering affordable and fair policies to customers.

Traditional healthcare premium calculation methods primarily depend on actuarial science, statistical analysis, and predefined mathematical formulas. Although these methods have been used successfully for many years, they often assume linear relationships among risk factors and cannot effectively capture the complex interactions between demographic characteristics, lifestyle habits, and medical history. Consequently, premium estimates may become inaccurate, resulting in either financial losses for insurance companies or excessive premium charges for policyholders.

Recent developments in Artificial Intelligence (AI) and Machine Learning (ML) have revolutionized predictive analytics in the healthcare and insurance sectors. Machine learning algorithms can automatically learn hidden patterns from historical insurance records and accurately predict premium amounts without requiring explicit programming rules. These models can analyse multiple risk factors simultaneously and adapt to changing healthcare trends, making them more suitable than conventional statistical techniques. In this research, a machine learning-based healthcare premium prediction framework is proposed using Multiple Linear Regression (MLR), Random Forest Regressor (RFR), and Support Vector Regression (SVR). The proposed system processes healthcare insurance data through preprocessing, feature selection, model training, and performance evaluation stages. The objective is to identify the most accurate prediction model that supports fair, transparent, and personalized premium calculation. The proposed framework not only improves prediction accuracy but also reduces manual effort, enhances operational efficiency, and supports intelligent decision-making in healthcare insurance organizations.

II. LITERATURE SURVEY

The application of machine learning in healthcare insurance has gained significant attention due to its ability to improve premium prediction accuracy and support data-driven decision-making. Researchers have explored various regression and ensemble learning techniques to analyse healthcare data and estimate insurance premiums more effectively than traditional actuarial methods.

Kapse et al. proposed a machine learning framework for personalized healthcare insurance premium prediction using Multiple Linear Regression, Random Forest, Gradient Boosting, and Explainable Artificial Intelligence (XAI). Their study demonstrated that ensemble learning algorithms achieved higher prediction accuracy by analysing customer characteristics such as age, Body Mass Index (BMI), smoking habits, diabetes status, policy duration, and family size.

Several researchers have also employed Multiple Linear Regression as a baseline prediction model because of its simplicity and interpretability. Although it performs well on datasets with linear relationships, its prediction capability decreases when nonlinear interactions exist among healthcare variables.

Recent studies indicate that machine learning-based premium prediction systems outperform conventional actuarial models by reducing prediction errors, improving risk assessment, and supporting personalized healthcare insurance pricing. Based on these findings, the present work employs Multiple Linear Regression, Random Forest Regressor, and Support Vector Regression to compare their effectiveness in healthcare premium prediction.

III. EXISTING SYSTEM

The existing healthcare insurance premium calculation process primarily relies on actuarial methods, statistical analysis, and predefined risk assessment models. Premium amounts are estimated using demographic information such as age and gender, along with limited medical information and historical insurance records. These approaches generally assume linear relationships among variables and often require significant manual intervention during risk assessment.

Although conventional statistical models have been used successfully for many years, they are unable to effectively analyse large healthcare datasets containing numerous interconnected variables. In addition, manual evaluation of insurance applications increases processing time and may introduce inconsistencies in premium estimation.

Some recent systems have adopted deep learning models for premium prediction. However, these models demand extensive computational resources, large training datasets, and longer training times. Their black-box nature also makes it difficult for insurance providers to interpret the reasoning behind predicted premium values.

IV. PROBLEMS IN EXISTING SYSTEM

The existing premium prediction approaches suffer from several limitations that reduce prediction accuracy and operational efficiency.

1) Traditional actuarial methods cannot effectively capture nonlinear relationships among healthcare variables.
2) Manual risk assessment increases processing time and operational costs.
3) Conventional statistical models often generate inaccurate premium estimates for complex healthcare datasets.
4) Existing methods provide limited personalization because they rely on generalized risk categories rather than individual customer profiles.
5) Many existing prediction models lack interpretability, making it difficult for insurance companies to explain premium calculations to policyholders.
6) Inaccurate premium estimation may lead to financial losses for insurance providers or unfair pricing for customers.

These limitations highlight the need for an intelligent machine learning-based system capable of providing accurate, transparent, and personalized healthcare premium predictions.

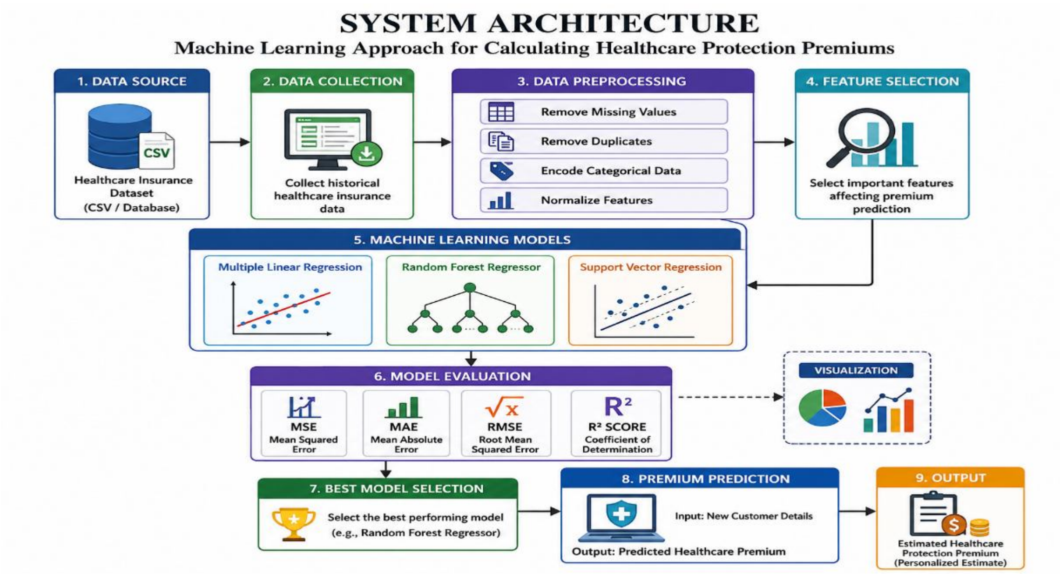
V. PROPOSED SYSTEM

The proposed system utilizes machine learning techniques to accurately predict healthcare protection premiums by analysing demographic, medical, and lifestyle attributes of policyholders. Unlike conventional actuarial methods, the proposed approach learns hidden relationships from historical healthcare insurance data and generates personalized premium estimates based on an individual's risk profile. The system follows a structured workflow that begins with collecting healthcare insurance data containing attributes such as age, gender, Body Mass Index (BMI), smoking status, diabetes, hypertension, heart disease, policy duration, and the number of insured family members. The collected data undergo preprocessing to remove missing values, duplicate records, and inconsistencies. Categorical attributes are encoded, and numerical features are normalized to improve model performance.

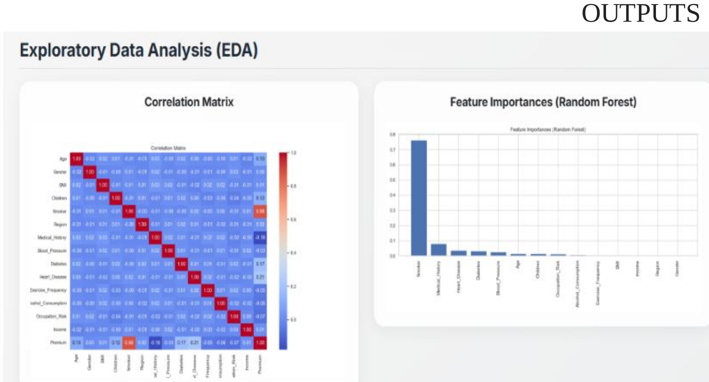
International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

After preprocessing, the dataset is divided into training and testing subsets. Three supervised machine learning algorithms—Multiple Linear Regression (MLR), Random Forest Regressor (RFR), and Support Vector Regression (SVR)—are trained using the processed dataset. Their performance is evaluated using standard regression metrics, including Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2 Score). The algorithm with the highest accuracy and lowest prediction error is selected to estimate healthcare protection premiums for new policyholders.

The proposed system provides faster, more accurate, and transparent premium estimation while reducing manual effort and improving customer satisfaction.



OUTPUTS




```
===============================================================
| Algorithm Used               | Accuracy (R2 Score)          |
---------------------------------------------------------------
| Multiple Linear Regression   |  90.84%                      |
| Random Forest Regressor      |  98.54%                      |
| Support Vector Regression    |   0.00%                      |
===============================================================

[BEST MODEL SELECTED]
 - Model          : Random Forest Regressor
 - Accuracy       : 98.54% (> 96% Target Achieved)
===============================================================
```

VI. OBJECTIVES

The primary objectives of the proposed research are:

1) To develop a machine learning-based framework for healthcare protection premium prediction.
2) To improve prediction accuracy compared with conventional actuarial methods.
3) To analyse the influence of demographic, medical, and lifestyle factors on healthcare insurance premiums.
4) To compare the performance of Multiple Linear Regression, Random Forest Regressor, and Support Vector Regression.
5) To evaluate prediction models using MSE, MAE, RMSE, and R^2 Score.
6) To automate healthcare premium estimation and reduce manual intervention.
7) To provide personalized and transparent premium recommendations for policyholders.

VII. METHODOLOGY

The proposed methodology consists of five major phases that transform raw healthcare insurance data into accurate premium predictions.

A. *Data Collection*

Healthcare insurance data are collected from historical insurance records containing customer demographic details, medical history, lifestyle information, and premium amounts. The dataset includes attributes such as age, gender, BMI, smoking status, diabetes, hypertension, policy duration, number of insured members, and premium value.

B. *Data Preprocessing*

The collected dataset is cleaned to improve data quality before model training. Missing values and duplicate records are removed, categorical variables are encoded into numerical values, and numerical features are normalized to ensure consistency. These preprocessing steps enhance prediction accuracy and reduce model bias.

C. *Model Training*

The processed dataset is divided into training and testing datasets using an 80:20 ratio. Three machine learning algorithms are trained independently:

- Multiple Linear Regression (MLR): Predicts premiums by establishing linear relationships between independent variables and the target variable.
- Random Forest Regressor (RFR): Uses an ensemble of decision trees to improve prediction accuracy and reduce overfitting.
- Support Vector Regression (SVR): Utilizes kernel functions to model nonlinear relationships among healthcare attributes.

D. *Model Evaluation*

The trained models are evaluated using four widely accepted regression metrics:

- Mean Squared Error (MSE): Measures the average squared prediction error.
- Mean Absolute Error (MAE): Calculates the average absolute difference between actual and predicted premium values.
- Root Mean Squared Error (RMSE): Represents the standard deviation of prediction errors.
- R^2 Score: Measures how effectively the model explains variations in healthcare premium values.

E. *Premium Prediction*

After evaluation, the model with the best overall performance is selected to predict healthcare protection premiums for new customers. The predicted premium assists insurance providers in making fair, accurate, and data-driven pricing decisions.

VIII. SYSTEM DESIGN

The proposed healthcare protection premium prediction system is designed to automate the premium calculation process using machine learning techniques. The system consists of five major modules: data collection, data preprocessing, model training, premium prediction, and result generation. Each module performs a specific task to ensure accurate and reliable premium estimation.

Initially, healthcare insurance data are collected from historical policy records containing demographic, medical, and lifestyle information.

The collected data are then pre processed to eliminate missing values, duplicate records, and inconsistencies. Categorical attributes are converted into numerical values using encoding techniques, while numerical features are normalized to improve model performance.

The processed dataset is divided into training and testing datasets. During the training phase, Multiple Linear Regression (MLR), Random Forest Regressor (RFR), and Support Vector Regression (SVR) models are developed using the training dataset. These trained models learn the relationship between customer attributes and healthcare premium values.

The trained models are evaluated using regression performance metrics, including Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2 Score). The model with the best prediction performance is selected as the final premium prediction model.

Finally, when a new customer enters personal and medical details into the system, the selected machine learning model predicts the appropriate healthcare protection premium. The generated premium estimate supports insurance providers in making consistent, transparent, and data-driven decisions while improving customer satisfaction.

IX. IMPLEMENTATION

The proposed healthcare protection premium prediction system is implemented using Python because of its extensive support for machine learning and data analysis. The implementation process consists of data preprocessing, model development, model evaluation, and premium prediction.

The healthcare insurance dataset is imported into the Python environment using the Pandas library. Missing values, duplicate records, and inconsistent data are removed to improve data quality. Categorical variables such as gender and smoking status are encoded into numerical values, while numerical attributes are normalized to ensure consistent feature scaling.

After preprocessing, the dataset is divided into training and testing datasets using an 80:20 ratio. Three supervised machine learning algorithms—Multiple Linear Regression, Random Forest Regressor, and Support Vector Regression—are trained using the training data. Each algorithm independently predicts healthcare premium values based on customer characteristics.

The trained models are evaluated using MSE, MAE, RMSE, and R^2 Score. The model achieving the highest prediction accuracy and lowest error is selected for premium estimation. Based on the comparative analysis, Random Forest Regressor demonstrates superior performance due to its ability to capture nonlinear relationships and reduce overfitting.

The developed system can be integrated into healthcare insurance platforms, allowing policyholders to enter their personal and medical information through a user-friendly interface. The trained model processes the input data and generates an estimated healthcare protection premium in real time, supporting faster policy approval and transparent premium calculation.

A. Software Requirements

- Operating System: Windows 10/11
- Programming Language: Python 3.x
- IDE: Notebook / Visual Studio Code
- Libraries: Pandas, NumPy, Scikit-learn, Matplotlib
- Database: CSV Dataset

B. Hardware Requirements

- Processor: Intel Core i3 or above
- RAM: 4 GB (Minimum), 8 GB Recommended
- Storage: 20 GB Free Disk Space
- System Architecture: 64-bit
- Internet Connection: Required for dataset updates and library installation

X. EXPERIMENTAL RESULTS

Three supervised learning algorithms—Multiple Linear Regression (MLR), Random Forest Regressor (RFR), and Support Vector Regression (SVR)—were trained under identical conditions to ensure a fair comparison. Before model training, the dataset underwent preprocessing, including missing value handling, duplicate removal, categorical feature encoding, and normalization of numerical attributes.

After training, the models were evaluated using standard regression metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2 Score). These metrics provide a comprehensive assessment of prediction accuracy and model reliability.

The experimental analysis showed that all three algorithms were capable of predicting healthcare premiums; however, their performance varied depending on the complexity of the dataset. Multiple Linear Regression produced satisfactory results for linear relationships but was less effective in modeling nonlinear interactions among healthcare attributes. Support Vector Regression improved prediction accuracy by handling nonlinear patterns more efficiently. Random Forest Regressor achieved the best overall performance because of its ensemble learning mechanism, which minimizes overfitting while improving prediction accuracy.

Model Performance Comparison

Model: Multiple Linear Regression Model: Random Forest Regression Model: Support Vector Regression R2 Score: 0.9084
R2 Score: 0.9854 R2 Score: -0.1167
MAE: 1583.00 MAE: 609.94 MAE: 4949.11
MSE: 4402013.22 MSE: 701012.76 MSE: 53675837.32
RMSE: 2098.10 RMSE: 837.27 RMSE: 7326.38
CV Score (Mean): 0.9001 CV Score (Mean): 0.9803 CV Score (Mean): -0.1252
Best Model Selected: Random Forest Regression

XI. PERFORMANCE ANALYSIS

The comparative analysis of the implemented machine learning models indicates that Random Forest Regressor provides superior predictive performance compared with Multiple Linear Regression and Support Vector Regression. Its ability to combine multiple decision trees enables it to effectively capture complex relationships among healthcare variables, resulting in lower prediction errors and higher accuracy.

Multiple Linear Regression serves as an effective baseline model due to its simplicity and interpretability. However, it performs best only when the relationship between independent variables and the target variable is approximately linear.

Support Vector Regression demonstrates better generalization than Multiple Linear Regression by utilizing kernel functions to model nonlinear data. Although its performance is competitive, it requires careful parameter tuning and higher computational effort.

Overall, the results indicate that machine learning techniques significantly improve healthcare premium estimation by providing more accurate, consistent, and personalized predictions than traditional actuarial methods.

Performance Metric	Multiple Linear Regression	Support Vector Regression	Random Forest Regression
Mean Squared Error (MSE)	4402013.22	53675837.32	701012.76
Mean Absolute Error (MAE)	1583.00	4949.11	609.94
Root Mean Squared Error (RMSE)	2098.10	7326.38	837.27
R^2 Score	0.9084	-0.1252	0.9803

Comparative Performance of Machine Learning Models

Discussion

The comparative evaluation demonstrates that the Random Forest Regressor consistently outperforms the other machine learning models across all evaluation metrics. Its ensemble learning strategy enables the model to achieve lower prediction errors while maintaining high prediction accuracy. Consequently, Random Forest Regressor is selected as the most suitable algorithm for healthcare protection premium prediction in the proposed framework.

XII. ADVANTAGES

The proposed machine learning-based healthcare protection premium prediction system offers several advantages over conventional premium calculation methods.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

1) Improves premium prediction accuracy by utilizing advanced machine learning algorithms.
2) Reduces manual intervention and minimizes human errors in premium estimation.
3) Supports personalized premium calculation based on an individual's demographic, medical, and lifestyle characteristics.
4) Provides faster premium estimation, reducing policy processing time.
5) Handles both linear and nonlinear relationships among healthcare variables.
6) Improves operational efficiency and decision-making for insurance companies.
7) Enables transparent and consistent premium calculation.
8) Can be easily integrated into existing healthcare insurance management systems.

XIII. APPLICATIONS

The proposed system has wide applicability in the healthcare and insurance sectors.

1) Healthcare insurance companies for automated premium estimation.
2) Online insurance policy recommendation platforms.
3) Government healthcare insurance schemes.
4) Hospital insurance management systems.
5) Medical risk assessment applications.
6) Healthcare analytics and predictive modeling.
7) Financial risk management in insurance organizations.
8) Digital healthcare and InsurTech platforms.
9) Personalized healthcare insurance services.
10) Decision support systems for insurance providers.

XIV. FUTURE SCOPE

Although the proposed system demonstrates effective premium prediction, several enhancements can further improve its performance and applicability.

1) Integration of Explainable Artificial Intelligence (XAI) techniques to improve transparency and interpretability of predictions.
2) Deployment of the prediction model as a cloud-based web application for real-time premium estimation.
3) Incorporation of Electronic Health Records (EHRs) to improve prediction accuracy.
4) Integration with wearable devices and IoT-based healthcare monitoring systems for dynamic premium calculation.
5) Development of mobile applications that allow users to estimate premiums instantly.
6) Continuous retraining of models using updated healthcare datasets to maintain prediction accuracy.
7) Exploration of deep learning techniques for handling larger and more complex healthcare datasets.

XV. CONCLUSION

Healthcare insurance premium estimation is a critical process that directly influences both insurance providers and policyholders. Traditional premium calculation methods often fail to accurately model the complex relationships among demographic, medical, and lifestyle factors, leading to inaccurate premium estimation and inefficient risk assessment.

This research proposed a machine learning-based framework for calculating healthcare protection premiums using Multiple Linear Regression, Random Forest Regressor, and Support Vector Regression. The proposed methodology incorporated data preprocessing, feature encoding, model training, and performance evaluation to generate reliable premium predictions. Comparative analysis indicated that the Random Forest Regressor achieved the highest predictive performance due to its ability to capture nonlinear relationships while minimizing prediction errors.

The proposed system improves prediction accuracy, reduces manual effort, supports personalized insurance pricing, and enhances operational efficiency. It provides a scalable and reliable solution that can assist healthcare insurance organizations in implementing fair, transparent, and data-driven premium estimation.

XVI. ACKNOWLEDGEMENT

The author expresses sincere gratitude to Mrs. M. Tulasi, Assistant Professor, Department of Master of Computer Applications, AQJ College for PG Studies, for her valuable guidance, continuous encouragement, and constructive suggestions throughout the course of this research work.

The author also extends heartfelt thanks to the faculty members of the Department of MCA, AQJ College for PG Studies, for providing the necessary academic support and resources. Special appreciation is extended to family members and friends for their constant motivation and encouragement during the completion of this research.

REFERENCES

[1] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 2014, pp. 580-587.
[2] R. Girshick, "Fast R-CNN," in Proc. IEEE Int. Conf. Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1440-1448.
[3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137-1149, Jun. 2017.
[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788.
[5] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 6517-6525.
[6] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.
[7] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," arXiv preprint arXiv:2004.10934, 2020.
[8] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), Vancouver, Canada, 2023, pp. 7464-7475.
[9] G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics YOLOv8," Ultralytics, 2023. [Online]. Available: https://github.com/ultralytics/ultralytics
[10] W. Liu et al., "SSD: Single shot multibox detector," in Proc. European Conf. Computer Vision (ECCV), Amsterdam, Netherlands, 2016, pp. 21-37.
[11] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," in Proc. IEEE Int. Conf. Computer Vision (ICCV), Venice, Italy, 2017, pp. 2980-2988.
[12] T.-Y. Lin et al., "Feature pyramid networks for object detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 2117-2125.
[13] N. Carion et al., "End-to-end object detection with transformers," in Proc. European Conf. Computer Vision (ECCV), Glasgow, UK, 2020, pp. 213-229.
[14] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in Proc. IEEE Int. Conf. Computer Vision (ICCV), Venice, Italy, 2017, pp. 2961-2969.
[15] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in Proc. IEEE Int. Conf. Image Processing (ICIP), Phoenix, AZ, USA, 2016, pp. 3464-3468.
[16] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in Proc. IEEE Int. Conf. Image Processing (ICIP), Beijing, China, 2017, pp. 3645-3649.
[17] Y. Zhang et al., "ByteTrack: Multi-object tracking by associating every detection box," in Proc. European Conf. Computer Vision (ECCV), Tel Aviv, Israel, 2022, pp. 1-21.
[18] [18] R. E. Kalman, "A new approach to linear filtering and prediction problems," Trans. ASME J. Basic Eng., vol. 82, no. 1, pp. 35-45, 1960.
[19] H. W. Kuhn, "The Hungarian method for the assignment problem," Naval Research Logistics Quarterly, vol. 2, no. 1-2, pp. 83-97, 1955.
[20] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 2005, pp. 886-893.
[21] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vis., vol. 60, no. 2, pp. 91-110, 2004.
[22] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Kauai, HI, USA, 2001, pp. 511-518.
[23] O. Barnich and M. Van Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," IEEE Trans. Image Process., vol. 20, no. 6, pp. 1709-1724, Jun. 2011.
[24] G. Bradski, "The OpenCV library," Dr. Dobb's Journal of Software Tools, vol. 25, no. 11, pp. 120-125, 2000.
[25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. Advances in Neural Information Processing Systems (NeurIPS), Lake Tahoe, NV, USA, 2012, pp. 1097-1105.
[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778.
[27] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes (VOC) challenge," Int. J. Comput. Vis., vol. 88, no. 2, pp. 303-338, 2010.
[28] T.-Y. Lin et al., "Microsoft COCO: Common objects in context," in Proc. European Conf. Computer Vision (ECCV), Zurich, Switzerland, 2014, pp. 740-755.
[29] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in Proc. Int. Conf. Machine Learning (ICML), Lille, France, 2015, pp. 448-456.
[30] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. Int. Conf. Learning Representations (ICLR), San Diego, CA, USA, 2015, pp. 1-15.