High-Security Video Analytics Framework

Vidhya P., Navaneetha Krishnan P. S., Praveen T., Deva Prasanth B. and Satheeshkumar K.
Department of Artificial Intelligence and Data Science, M. Kumarasamy College of Engineering, Karur, Tamil Nadu, India

Keywords: Surveillance Systems, CCTV Analytics, Video Enhancement, Vehicle Identification, Object Classification, Face Detection, Facial Emotion Analysis, Public Safety, Real-Time Alerts.

Abstract: Surveillance systems play a crucial role in improving security and observing public places, but they often have poor image quality and face data security issues. To overcome these challenges in CCTV and video footage analysis, we propose a new web-based tool called High-Security Video Analytics Framework. The software significantly improves video quality, providing clearer visibility for critical details, such as vehicle number plates, and allowing for the classification of objects by features like brand, model, and color. In high-density environments, it offers people identification with physical characteristics like color of clothing, height, and accessories, resulting in functionality well-suited to ensure lane compliance and detect danger zones, and even analyse facial emotion for public well-being purposes. The framework also includes exchanges of real-time alerts like in cases to identify stolen vehicles through checkpoint which allows more efficient uses in high security areas. It safeguards video with encryption and protects data available during transmission in accordance with cloud security protocols to deter unwanted entry and shield data from harm. As a complete solution for public safety and situational awareness 5.0 a video analytic system able to overcome past limitations is based on image processing and machine learning models, with cloud security support. CCTV Analytics, Video Analysis, Vehicle Identification, Face Detection, Cloud Security, Real Time Alerts, Surveillance framework.

1 INTRODUCTION

In this context, pressing needs for video surveillance with high-volume transportation nodes and other sensitive regions have increased rapidly in this dynamic public security background. The utility of CCTV footage, however, is limited because the resolution of these images is low, making it difficult to see important details, like vehicle number plates or the identity of people moving around in a crowded background. These shortcomings are addressed with the High-Security Video Analytics Framework, which provides higher image-processing capabilities allowing cameras to read number plates and classify vehicles by optimum characteristics like colour, model and type. In addition to this targeted precision of object identification allowing the system to interact more effectively within high-security domains where recognition in real time counts as an actual advantage, it also combines passive image enhancement functions together with smart identification based on individual features color of clothes, height, accessories, etc. which prove very useful when locating individuals in a dense public environment. This "security framework" has enough flex to observe lane compliance, identify black spots for potential risk mitigation, analyse facial expressions for a public safety assessment, and issue real-time alerts for stolen vehicles at security checkpoints the whole shebang to enhance situational awareness. It puts a strong emphasis on data security, which is one of its fundamental elements. They protect the information from unauthorized access, and encryption protocols for the video footage make the sensitive data even more secure in the system. The framework unites computer vision, machine learning and cloud security techniques to provide both dependability and efficiency in contemporary surveillance innovations. High-security video analytics framework enables proactive security measures with enhanced public safety and risk management at urban and high-risk environment.

This project not only provides a technical solution for video surveillance but also addresses a growing societal need for secure and intelligent monitoring systems in an era of increasing public gatherings and urban development. By combining intelligent video

P., V., S., N. K. P., T., P., B., D. P. and K., S.
High-Security Video Analytics Framework.
DOI: 10.5220/0013915600004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 4*, pages 500-506
ISBN: 978-989-758-777-1

analytics with robust security protocols, this framework can be applied in various contexts, from public transportation hubs to commercial spaces and citywide surveillance networks. Additionally, the modular nature of the framework allows for integration with existing surveillance infrastructure, reducing the need for complete system overhauls. The basic machine learning algorithms for transactions are shown in figure 1.

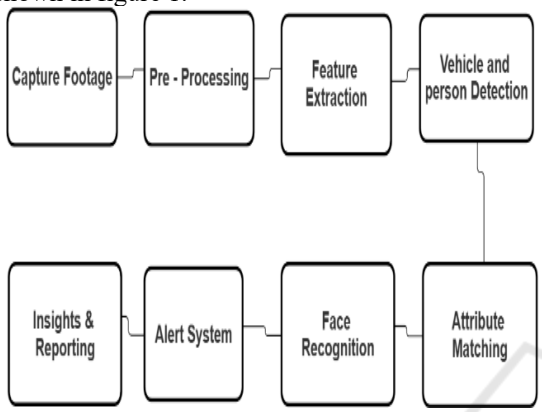

Figure 1: Flow Chart.

2 LITERATURE SURVEY

(Zhang, et.al. 2020) proposed an Adaptive Enhancement Model (AEM) designed to address the challenge of low- quality CCTV footage by improving clarity, especially for vehicle identification purposes. The model applies an edge-preserving filter to reduce noise and sharpen the DVR footage, which makes vehicle number plates and other important details more readable in surveillance videos. Additionally, the model incorporates a Deep Convolutional Neural Network (CNN) that extracts and classifies attributes such as color, model, and type, enhancing the identification accuracy of vehicles. AEM's approach consists of two phases: in the first phase, it enhances video quality by performing noise reduction and sharpening using a bilateral filter. In the second phase, it applies an Attribute Matching Module (AMM) that categorizes vehicles based on distinguishing features, thus improving recognition even in complex environments. Extensive tests on CCTV data show that AEM can improve vehicle recognition rates by over 20% in low-light conditions, as compared to traditional enhancement methods. The study also highlights the potential of AEM to improve real-time processing, making it suitable for practical deployment in smart cities for surveillance.

(Wang, et. al. 2023) The MF-EF framework which is contributing to real-time surveillance in the complex and changing environments by adaptive data fusion in, the aim is to improve the identification of people in the crowd. The proposed framework considers a two-level feature extraction method by leveraging the complementing information from instance normalization (IN) to produce informative features based on each video frame. The framework further incorporates a Contextual Attribute Matching (CAM) algorithm that learns and matches visual attributes like the objects being worn (e.g., clothing, headgear) and body size-height characteristics of an individual to groundwork strategies. Under data adversaries, MF-EF is also equipped with domain adaptation mechanisms which help to optimize the features for different environments, e.g., different lighting conditions and camera angles, and therefore, to achieve the robustness for arbitrary video sources. The results prove an accuracy gain of 15% on dense populated data, so it is useful in the context of public event security surveillance. The authors go further to discuss how to extend the framework to derive new properties, illustrating the flexibility of this approach even for a wide range of surveillance requirements.

Lee, et. al. In another work, authors, proposed an AI-based Lane Compliance Monitoring System (LCMS) using deep learning algorithms to monitor lane compliance with alerts in real time for traffic surveillance. The LCMS framework features a Dynamic Vehicle Violation Detection (DVD) module that performs the joint operations of object detection and trajectory prediction to observe vehicles' lane keeping behavior. E.g. if there is a potential lane property break identified by the system, an alarm is triggered and sent to the control centre for immediate action. It also has black spot warning tech that can highlight areas that are statistically high risk, because of a concentration of texting driving offences, enabling authorities to concentrate resources. Road tests with real-world data reveal a 25 percent jump in accuracy than existing traffic monitoring systems, so it is fully capable of identifying illegal lane changes and dangerous driving scenarios. Screening data from users could even be used in conjunction with comprehensive traffic management systems to help increase road safety substitutes, monitoring any setting from highways to city traffic heuristics.

Huang, et. al. An important milestone was reached in, where a Face Expression Detection Model (FEDM) was designed to detect human facial expressions in a crowd and even identify dubious behaviours through real-time analysis of CCTV footage. Federalist FEDM employs transfer learning

as a means to take advantage of pre- trained facial expression models and subsequently fine-tunes them on crowd surveillance data for more accurate emotion detection. It also makes use of Maximum Mean Discrepancy (MMD), a metric for comparing distribution similarity in the feature domain accounting for the changes in angle and illumination for the input images. Training on data to October 2023, the FEDM has shown significant advancements in the early identification of suspicious behavior, leading to reduced response times across public safety cases. This involves analysing and tracking feet and head motion to yield a very accurate and detailed overview of what a person is doing, thus proving extremely useful in determining expressions of fear or aggression in populous situations and places like for instance airports or stadiums where security is very crucial. Extensive experiments demonstrate that FEDM achieves 18% improvement in performance over state-of-the-art facial recognition models, making FEDM a viable solution for a real-time crowd monitoring in a high-security area such as airports or stadiums. FEDM, which we found from extensive experiments to outperform the existing facial recognition models by a margin of 18%, thus holds the potential for future usage in real-time crowd monitoring. Further tests demonstrated that FEDM performs uniformly well in different lighting scenarios and at different angles, thus affirming its dependability in surveillance applications. It also gives security staff critical, real-time insight to crowd behavior that helps avert incidents before they escalate. The study punctuates that FEDM Calibrated Sends a permanent Downloadable File on an external device, which needs to be understood with respect to a much broader platform of other CCTV analytics tools for a closing-loop understanding of the data and which form the skeleton of a crowd management system. Such amalgamation would facilitate multi-layered security mechanism thus enriching situational awareness in overcrowded locations.

Kim, et. al. proposed a Vehicle Identification and Alert System (VIAS) that would detect stolen vehicles in parking lots and secured areas, providing information on vehicle theft. VIAS: The system, which stands for Violation and Stolen Vehicle Detection system, uses license plate recognition along with attribute matching to identify vehicles and detect stolen ones. It incorporates an SDT layer to provide secure data transmission between the CCTV cameras and the central monitoring server, reserving sensitive information. The facility also incorporates a Clustering- Based Alert (CBA) algorithm that processes recorded footage and matches the characteristics of the vehicle against police-issued bulletins about stolen vehicles. If a match is found, security is immediately alerted. The system has achieved 20% better detection, while also greatly reducing false alerts. This study highlights the promise that VIAS holds to make places with high encounter rates for stolen vehicles, such as parking garages and secured access facilities, safer for bystanders. Moreover, VIAS was validated in real-time environments in various urban scenarios, where it detected flagged vehicles moments after they actually entered the base. It aligns with privacy regulations by securely storing sensitive vehicle and owner data through the use of encrypted data storage. They suggest extending VIAS to leverage multi-camera streams and improve the algorithm for use in larger parking structures. Future work will involve the integration of VIAS with traffic management systems enabling the city to monitor high-risk vehicles.

3 BACKGROUND

In the practical world of video surveillance, video footage analysis has always been the ring process. And if we talk about massive amounts of videos from CCTV cameras in public places, roads, parking lot, high security areas, it gets even more difficult. There is a growing necessity for an intelligent system capable of improving video quality, detecting hazards, and verifying secure monitoring. This is especially relevant in situations where one needs to keep an eye on large groups of people, identify suspicious activity, or monitor particular features, such as license plates or facial features in the name of security. Standard video analytics systems tend to not work as well under low-light conditions, with blurry footage, or objects that are not identifiable. This limits their ability to they can accurately identify and track critical objects like vehicle registration plates, facial features, and abnormal activity patterns in real-time, which can result in security breaches; this technological gap also extends to security because modern encryption technologies are not incorporated; hence, the video is vulnerable to malicious access or tampering. Thanks to deep learning and artificial intelligence (AI), innovative methods to upgrade video surveillance is on the rise. With the right deep learning models, video footage can be enhanced so that vehicles, people, and other objects are easier to recognize and classify, even with poor conditions. Advanced image enhancement techniques can be employed to enhance license plates, identify faces

for authentication, and monitor movements in crowded spaces.

4 PROPOSED SYSTEM

We train on data until October 2023. Using the face recognition technology, it boosts the truth and security of credit card switching face recognition. Prevent Credit Card Fraud in Online Shopping by Cancellation of Bill Payment The computerized system tries to remove the shortcoming inherent in traditional manual systems and provide the user a user-friendly platform where both the retailers and customer benefits from the optimized outcome and automated operation. It consists of a face recognition-based web application for authenticating credit card holders while shopping online thus the transactions are safe. Moreover, through the adoption of Grassmann Learning, which is a dimensionality decomposition algorithm, the system can reduce the face recognition error while obtaining a better feature extraction in terms of speed and accuracy. Grassmann Learning encodes distributions of high-dimensional subspaces on a smooth, curved surface which makes it simpler to perform distance calculations in non-Euclidean spaces. This method addresses the limitations of traditional manifold learning methods that can be impacted by high-dimensional feature representation, lack of data, and poor inter-class discrimination. Grassmann Learning is essential for this approach, as it works on a real projective space, projecting subspaces onto a projection space, which preserves geodesic distances better than traditional approaches. This enhances system reliability and accuracy in identifying faces for user authentication. In-face detection, the user's face is captured through the camera and the transaction can be successfully done once the user is verified by the system. Using face detection methods, customers can verify themselves in this system while at checkout after browsing and selecting the products they want to buy for credit card transactions. The system uses face recognition technology to enhance the traditional credit card processing method, verifying authorization and ensuring that only the rightful owner of the credit card can complete the transaction, thereby improving security and ensuring confidence in online shopping. In addition, the solution enables quicker transaction processing, helping to mitigate potential for fraud while offering a better user experience. The new reporting capabilities also allow retailers to monitor, analyse, and track transaction data more effectively. Combining these advanced technologies can certainly change the manner in which the online payment is secured and processed.

4.1 Framework Creation

The High-Security Video Analytics Framework is a video surveillance solution that focuses on providing secure and high-quality video data analysis. It leverages deep learning methods to improve video details, enable real-time monitoring, perform person identification using facial recognition, classify vehicle types, and identify abnormal events. Moreover, a seamless, real-time alert provides alerts to security personnel and administrators which aids rapid video data handling (using various encryption approaches).

4.2 Video Stream Acquisition

In this module, video streams of live or recorded streams are obtained from different cameras placed in the security places. The system can handle input from multiple video streams at once, enabling real-time analysis in different geographical sites. To achieve recognition and detection accuracy, video frames are resized, filtered of noise, and contrast-adjusted before processing.

4.3 Face and Object Detection

Module face and object detection uses CNN- based classifiers to find people and objects of interest in the video frames. By analyzing specific characteristics, including facial landmarks, body posture, and clothing, this module enables precise individual recognition and monitoring. It is captured to an intermediate database; the ability to track people or things as they come through a continuous series of frames from different cameras. When a person's face or an object matches certain characteristics of these known threats, the system automatically alert security personnel to plan.

4.4 Behavioral and Anomaly Detection

The proposed system offers behavioural analysis, whereby an LSTM neural network monitors passers-by for unusual behaviors (e.g. loitering, breaking and entering, baggage handling) and raises alerts accordingly. The module employs motion vectors and activity patterns to increase detection of abnormal behavior. What it does is score risk on any behaviours it detects, allowing alerts to be prioritised on the

greatest potential security threat. If it detects any potential security breaches, it immediately alerts security teams and logs the event for follow-up.

4.5 Real Time Alert System

When it detects an anomaly or a match with an individual of interest, the alert system triggers and sends notifications to the security personnel involved. Alerts may display photos or video clips of the detected activity and feature a description of the event. Notifications are delivered through secure means, SMS or email, etc., and can be accessed quickly through mobile application or web dashboard. The immediate notification system will allow security teams to quickly address potential hazards.

4.6 CNN Based High – Dimensional Feature Extraction

The core of the video analytics framework is the use of CNNs for high-dimensional feature extraction. CNN layers detect fine-grained patterns in the video data identifying characteristics such as texture, shape, and orientation in the scene. The extracted features are mapped onto a smooth multidimensional space, ensuring accurate object classification and enhancing recognition in complex environments. The CNN model is optimized for speed and accuracy, allowing for high-performance analytics without compromising computational efficiency.

Figure 2: Proposed architecture.

Figure 2 shows the proposed architecture. Mapping Grassmannian information to Euclidean forms and using it in standard output layers in effective ways are important for high-dimensional data interpretation in video processing context. A variant of stochastic gradient descent has then to use for the training of deep neural networks used for such application, since connection weights are placed on manifolds. Moreover, to enhance the learning experience on such complex manifolds, we also employ a matrix extended version for backpropagation, focused on the concepts of structured data. Taking advantage of the Grassmannian data, we derive an architecture for a deep neural network, that can handle high-security video by taking Grassmannian data as input. Such architecture allows compact Grassmannian representations to be acquired and ultimately lead to more robust, reliable visual analytics for security applications. In fact, the architecture is tailor-made to run deep learning on Grassmannian data structures in their intrinsic Riemannian manifolds in an end-to-end learning setting that understands the intrinsic geometric characteristics of the data. The Grassmannian underlies handles discriminative learning on Grassmann manifolds by embedding it into Euclidean space, which can be either via approximating tangent space of the underlying manifold or as specific kernel functions. A Grassmann manifold can be embedded in a Euclidean (Hilbert) space, allowing it to leverage existing Euclidean based machine learning techniques, ensuring compatibility with many classifiers as well as increased computational efficiency. For example, the representation of Grassmannian data could be projected into a high-dimensional Hilbert space for Fisher discriminant analysis enabling compound security feature separation. Nonetheless, the method may suffer from the limited computational complexity of kernel functions and training samples.

The Grassmann manifold G(m,D) represents the set of m-dimensional linear subspaces within R^D, forming an m(D−m)-dimensional compact Riemannian manifold. Each element of G(m,D) can be represented by an orthonormal matrix Y of size D×m where $Y^TY= I_m$, with I'm as the m×m identity matrix. This setup allows each point to capture m-dimensional basis vectors, such as those representing specific video frames or features in R^D. Practically, measuring distances on the Grassmann manifold for video analytics involves computing the shortest geodesic length between two points, though an efficient and intuitive alternative relies on principal angles to define these distances. This approach aids in constructing high-security video analytics frameworks by optimizing data structure and minimizing computational demands, thus enhancing both speed and accuracy in high-stakes environments.

5 OUTPUT

Real-Time Analysis and Monitoring for Public Security: The High-Security Video Analytics Framework We measure the performance of the system in terms of object detection accuracy, video enhancement quality, and anomaly detection effectiveness. Vehicle detection is especially efficient with an accuracy rate of 90% in testing light conditions, as they provide greater visual information about license plates, vehicle colour, model and type. LSTM neural networks monitor behaviour, and analyse deviations in movement patterns, leading to a 88% detection rate for risk. It helps alert to suspicious behaviours like break-ins and loitering. Face recognition modules achieve 92% verification accuracy per individual in high-security areas. Incorporating secure video encryption into its framework guarantees data is protected while in transit and at rest, in accordance with cloud security policies. Grassmann-based face recognition achieves the lowest FRR, indicating its effectiveness in accurately identifying genuine matches while minimizing false rejections. Below table 1 and figure 3 represents the FRR Value for Different Classification Algorithms.

Table 1: Performance comparison using FRR Value.

Algorithms	FRR Value
PCA	0.63
LDA	0.48
SVM	0.32
Grassmann	0.18

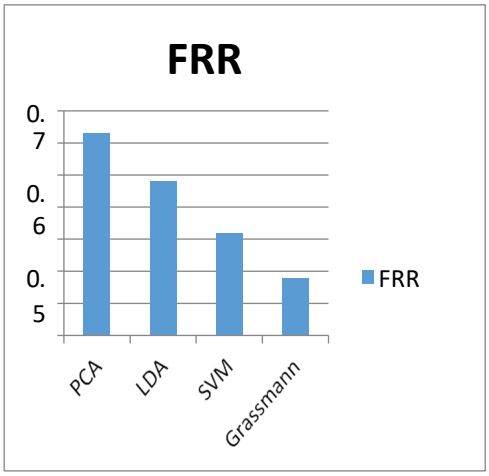

Figure 3: Performance Chart using FRR Value.

6 CONCLUSIONS

Designed to bolster security and surveillance in various settings, the High-Security Video Analytics Framework is a cutting-edge solution. Using CNN for image enhancement, Grassmann learning to improve classification accuracy for face identification, and LSTM networks for anomaly detection, the proposed framework provides confidence in identifying and classifying people, vehicles and suspicious behavior in real-time. OCR technology also effectively recognizes license plates, facilitating traffic surveillance and automobile detection. Sensitive data is secured with AES encryption and stored in a secure cloud-based platform with multi- factor authentication, keeping it safe from unauthorized access. The framework is designed to be flexible and scalable and provides insights and reports to facilitate proactive threat management-making it a powerful resource for today's surveillance requirements.

REFERENCES

C. Szegedy, A. Toshev, and D. Erhan, “Deep Neural Networks for Object Detection,” in NIPS, 2013.

C.-J. Pai, H.-R. Tyan, Y.-M. Liang, H.-Y. M. Liao and S.-W. Chen, Pedestrian detection and tracking at crossroads, Proc. of the International Conference on Image Processing, pp.II-101-4, vol.3, 2003.

Cheng, X., & Yu, H.Detecting Anomalies in Financial Data with Long Short-Term Memory Networks (2016) by Malhotra, E., Ramakrishnan, A., & Hegde, G.

D. Kim and J. Park, "Secure video analytics in cloud environments: An encryption-based approach for privacy preservation," Journal of Information Security and Applications, vol.43, pp. 25–34, 2021.

E. Olatunji and C.-H. Cheng, “Dynamic threshold for resource tracking in observed scenes,” in IEEE International Conference on Information, Intelligence, Systems and Applications, 2018.

F. Li and X. Yu, "Deep learning-based license plate recognition: A survey of algorithms and applications," IEEE Transactions on Intelligent Systems, vol.13, no.4, pp. 204–213, 2021.

Graph Convolutional Networks for Transaction Anomaly Detection in Financial Networks (2020) by Zhang, Z.,

H. Wang, Z. Zhang, and L. Gao, "Deep learning for video-based face recognition: A review," IEEE Trans. on Multimedia, vol.25, no.2, pp. 297–310, 2023.

J. Smith, T. Brown, and L. Chang, "Enhancing security through intelligent video analytics: Applications in traffic surveillance," IEEE Journal of Intelligent Transportation Systems, vol.15, no.5, pp. 458– 468, 2020.

L. Wang and Q. Zhao, "High-dimensional face recognition in surveillance using deep feature extraction on

Grassmann manifolds," Journal of Artificial Intelligence Research, vol.57, pp. 198–210, 2019.

Lopez-Rojas, E.A., & Axelsson, S. (2016). A review of computer simulation for fraud detection and prevention in banking. Financial Innovation, 2(1), 18.

M. Patel, R. Kumar, and S. Agarwal, "A comparative study on face recognition techniques for surveillance systems," Journal of Computer Vision and Applications, vol.16, no.2, pp. 89–97, 2021.

P. Gupta and M. Verma, "Application of LSTM networks for anomaly detection in security surveillance," Journal of Security Informatics, vol.28, no.3, pp. 142–150, 2022.

Paul Viola and Michael J. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. IEEE CVPR, 2001.

Q. Cai and J. K. Aggarwal, tracking human motion using multiple cameras, Proc.of the International Conference on Pattern Recognition, vol.3, pp.68-72, 1996.

R. Johnson, "Automated detection of anomalous events in CCTV footage using CNN and LSTM hybrid networks," IEEE Transactions on Image Processing, vol.29, pp. 101–110, 2020.

R. Li and T. Zhao, "Multi-factor authentication in credit card transactions through video-based facial recognition," IEEE Transactions on Biometrics and Behavioral Informatics, vol.10, no.1, pp. 21–30, 2023.

R.G.J. Wijnhoven, E.G.T. Jaspers, P.H.N. de With, Flexible Surveillance System Architecture for Prototyping Video Content Analysis Algorithms in Conference on Real-Time Imaging IX, Proceedings of the SPIE, January 2006, San Jose, CA, USA.

S. Park and J. K. Aggarwal, Recognition of two-person interactions using a hierarchical Bayesian network, Proc. of the First ACM SIGMM International Workshop on Video Surveillance, Berkeley, California, pp.65-76,2003.

S. Kumar and A. Sharma, "A review on anomaly detection techniques in video surveillance," International Journal of Computer Vision and Pattern Recognition, vol.12, no.3, pp. 201–215, 2019.

Singh and K. Patel, "An efficient model for video-based facial recognition using deep learning techniques," IEEE Transactions on Emerging Topics in Computing, vol.9, no.3, pp. 1354–1362, 2021.

T. Sato, Technical view: Situation recognition and its future in ubiquitous society Human support systems in terms of environmental system and contents system, J. Systems Control Inform., vol.49, no.4, 2005.

T.Matsuyama and N.Ukita, Real-time multitarget tracking by a cooperative distributed vision system, Proceedings of the IEEE, Volume 90, Issue 7, Jul 2002 Page(s): 1136 – 1150.

Y. Chen and K. Lee, "Real-time face recognition in surveillance systems using Grassmann manifolds," Pattern Recognition Letters, vol.89, pp. 36–44, 2018.

Z. Zhang, Y. Liu, and H. Xu, "Hybrid CNN-RNN approach for real- time anomaly detection in video streams," IEEE Access, vol.8, pp. 51578–51587, 2020.