International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 15s, 2025
https://www.theaspd.com/ijes.php

Optimized Image-Based Event Detection Via K-Fold Learning And Minimum-Boundary Deep Cnns

S. Sai Satyanarayana Reddy[1*], Seelam Meghana Reddy[2], Seelam Sanjana[3], Angadi Seshappa[4]
[1]Department of Computer Sciences and Engineering, KG Reddy College of Engineering &Technology, Hyderabad, TS, India, saisn90@gmail.com
[2]Technical consultant, Darwinbox, Hyderabad, TS, India, ssmegu900@gmail.com
[3]Lead Application support Analyst Geogetown University, McDonough School of Business, USA, anjanaseelam75@gmail.com
[4]Department of Mechanical Engineering, KG Reddy College of Engineering &Technology, Hyderabad, TS, India, seshurambollimi@gmail.com
*Corresponding Author: saisn90@gmail.com

Abstract– *Detecting events from real-world photographs and videos is a significant and intricate endeavour. Historically, forecasts about traffic, population dynamics, and environmental hazards mostly depend on sensor data rather than direct picture analysis. Images are first transformed into binary data derived from photographs or video sources. Despite extensive research on event identification, current methodologies often encounter accuracy issues stemming from inaccurate social network data, asymmetric datasets, restricted application scope, and inadequate semantic feature extraction. This paper presents a Minimum Boundary Deep Convolutional Neural Network (MB-deep CNN) to overcome these restrictions. The model improves performance by extracting information from picture intensity, directional patterns, gradient patterns, and spatial properties, enabling efficient pixel translation. The MB-deep CNN segments and recognises events directly from visual inputs, enhancing accuracy and beyond the limitations of prior approaches.*
Keywords– *Convolutional Neural Networks, Deep Learning, Event Detection, Image Analysis, Minimum Boundary Hyper plane, Pattern Recognition*

I. INTRODUCTION

Protection monitoring, analysing footage, and improving interaction between users and computers are just a few of the many areas where automated recognition of events has shown to be an invaluable tool[5]. As video-based applications continue to expand, event detection plays a vital role in filtering, searching, indexing, and mining large video datasets[6]. In opposition, implementing these systems on stationary platforms, mobile vehicles, or UAVs remains challenging, particularly in dynamic settings such as urban environments and battlefields, where extracting crucial information from video streams is essential[2]. Events captured in video can range from subtle action slick facial expressions, gestures, and postures to complex, large-scale interactions between moving entities and static scene elements[9]. Integrating advanced techniques into event detection methods has improved the automated recognition of lower-level semantic patterns, including shot boundaries, semantic units, and event classification[3],[11-14]. Accurate identification of individual with transport mobility is crucial to video surveillance and monitoring (VSAM), which is often centred around scene activities. By concentrating on pictures that elicit substantial consumer attention, these innovations not only save processing time and expenses, but they also minimise human intervention. To further improve identification precision, optical characteristics throughout the scene are spatially correlated[2-3].

In order to detect probably dangerous actions, it is necessary to know when and how they occurred. Whether you're following a smaller group of people or a bigger one, activity recognition is crucial in complicated situations [4]. By successfully modelling and recognising simple events, such as intrusion detection, techniques like Hidden Markov Models (HMMs) have shown their versatility in event recognition tasks[8, 15]. New approaches recast phenomenon recognition as a simple case of categorical identification. Nevertheless, there are still obstacles to effectively detecting anomalies in video material [1]. Although common approaches may effectively analyse activity by transforming pixel data into lower-

level characteristics, reliably detecting complicated, protracted events remains a challenge. Algorithms as they stand can assess basic human activities, but they often struggle in noisy contexts [5].
Instruments such as overall Harrison edge analyser have improved the efficiency of image identification using visual recordings that include motion. This paper presents the Least Border Deep Convolutional Neural Network (MB-deep CNN) as a new tool for improving image-based event identification. From photos fed into it from real-time datasets, this model extracts a wide range of properties, including magnitude, orientation, transition structures, deviation, even spectrum centring. After that, it uses an SVM (Support Vector Machine) and an Absolute Barrier hyper plane to map information towards a higher-dimensional space in order to identify while categorise activities. Optimising segmentation procedures, lowering bias, and improving the efficiency of feature selection and event detection are all achieved by the MB-deep CNN by decreasing feature dimensionality[1],[16-23].

II. LITERATURE REVIEW

The use of convolutional neural networks (CNNs) has improved an ML-based method for anomaly detection in visual and audio recordings. In order to reduce distortion in the data, data pre-processing methods have been implemented preceding obtaining features. The classifier was able to successfully identify events in both still photos and moving ones, however predicting events in recordings required a lot of computer computational resources and time [1]. Using multi-layered convolutional neural networks (CNNs), a deep learning (DL) method was also suggested for precisely localising and timing events [2]. But even after going through many rounds of dataset refining, the approach still had poor localisation accuracy in complicated contexts since there wasn't enough training information. Auditory pattern analysis was the centre of another approach that aimed to identify events [3]. The system used Hidden Markov Models (HMMs) to arrange subcategories according to keywords, preprocess audio input, and classify patterns. The method worked well, but it was computationally expensive and hard to maintain, particularly when dealing with dense datasets that had dynamic properties. A graph-based approach to event detection used sensors to gather data from the actual environment, data proved then transformed into pictures for the purpose of obtaining features. Although this approach successfully identified occurrences, it exhibited difficulty adapting to new occurrences and accurately localising individual objects [4]. Automatic analysis of the form and activity of human motion was carried out using temporal logic networks in a visual-based method for behaviour recognition and movement tracking. Despite its usefulness in enhancing our comprehension of interpersonal behaviours, such method struggled to effectively encode shorter information along with deal with complicated, high-level tasks [5].
To decipher video-based occurrences, a subspace-based audiovisual processing approach included a Decision Tree (DT) component. The analytic method became more complicated as a result of the greater data quantities that resulted from this technique [6]. Additionally, CNNs were used to differentiate between typical and out-of-the-ordinary occurrences in videos. Despite drawing on previous methods, this strategy had initial difficulties in accurately detecting anomalies and required a large amount of time for data collecting, pre-processing, and feature extraction [7]. Lastly, events were classified using the size and form of objects in photos or videos using Hidden Markov Models (HMMs). While this strategy was effective in identifying events, it did not thoroughly identify all abnormalities [8].

III. PROPOSED METHODOLOGY

The objectives of the project include creating and implementing an occurrence identification system that uses images and incorporating the suggested MB-Deep CNN classifier. It all starts with pre-processing, when images taken within the Human collection [36] are cleaned and afterwards made to seem better. These next step is to separate attributes from the pictures, such as magnitude, orientation, range, other spectrum information, in order to find formations across them. Indices aid in defining boundaries as well as adjusting picture addition, meanwhile directions aid in integrating programming data. Spectrum structures compute features like spectroscopic median as skewing, whereas gradation structures concentrate on identifying picture variances based on gradients. Certain spectrum characteristics are useful for analysing symmetry, orientation coordination, picture magnitude, especially spectroscopic a centroid, these stand for the visual centre of mass of the photograph.

After these characteristics are retrieved, they are combined and sent into the MB-Deep CNN classifier to produce a feature set. This classifier does double duty by reducing the likelihood of bias in the classification process and guaranteeing accurate identification of events. Fig. 1 is a graphical representation of the recommended model's entire process.

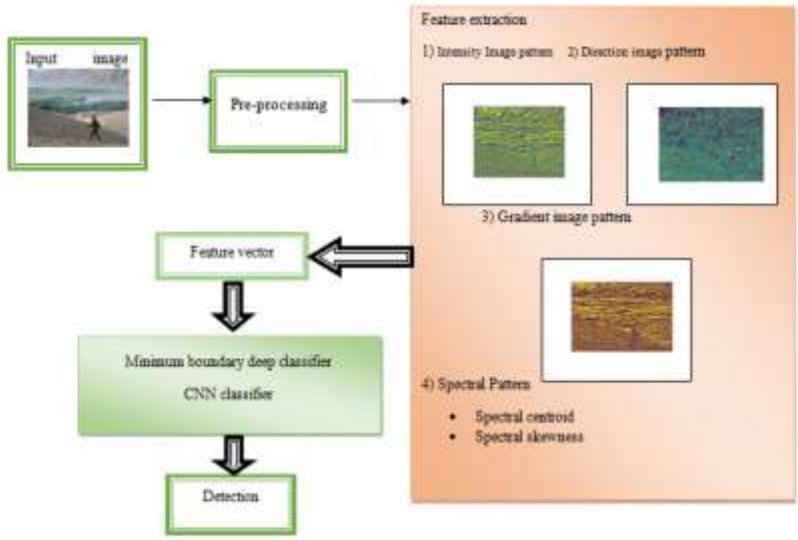

Fig. 1 This is the MB-Deep CNN classifier's circuit topology.

A.MB-Deep CNN Classifier For Event Detection: A photograph categorisation that activity identification machine learning encoder called the MB-Deep CNN is presented in the present article. To guarantee precise occurrence identification, the algorithm is designed to collect multidimensional features. The design of this connection is a completely interconnected cortical networks that culminates in matched convolution layers (CL) with layers that maximally pool [31, 32]. These paired layers interact through activation functions to perform feature extraction, dimensionality reduction, and image segmentation.

The MB-Deep CNN employs deep convolutional neural networks (DCNNs) to establish connections between similar neurons in images while optimizing weight and bias parameters [29]. As illustrated in Fig. 2, the model processes input data through 2D convolution layers, structuring it into pixel values. The data is then reorganized based on these values and batch-normalized for standardization. Subsequently, max-pooling is applied to extract the highest pixel values within localized regions of the image.

$$a_I^M = l\left(\sum_{b=1}^{map} a_{f_{vc}}^{M-1} * S_{f_{vc}}^M + D_{f_{vc}}^M\right) \qquad (1)$$

A convolutional regions may be identified using the formula provided, somewhere a_I^O stand for this preliminary facet chart of $M-1$this deposit, also a_I^M be this $(I)-th$deposit with mapcommunicates those numeral with contribution charts, S_{IB}^M also D_I^M are the restrictions within CNN through experimentation, with $l(\cdot)$its employed for states that non-linear utility, with $*$ represents for states its procedure with involvedness.S_{out} its number of parameter standards, while the number of overlapping stages is represented by O.

The hyper plane in a Support Vector Machine (SVM) is determined by the equation: $w \times x + b = 0$, where: ***w*** is the **normal vector** to the hyper plane, ***b*** represents the **offset (bias)**, and ***x*** denotes the input feature values.

An SVM can be trained using feature vectors derived from extracted features. The algorithm constructs a hyper plane in high-dimensional space (often projected into fewer dimensions) to separate data into distinct classes. This separation allows the SVM to classify data points, even in a two-dimensional feature space, by optimizing the decision boundary based on weight (w) and bias (b).

The classification process relies on binary decision-making, where the optimal hyper plane is selected to maximize the margin between classes. The model's performance depends on minimizing classification errors while maintaining the maximum-margin separation.

$$biclafun = SIG\left(\sum_{a=1}^{0} \partial_{a_1^M} \ker nal(E) - \lambda\right) \quad (2)$$

anywhere E its instruct effort structure to undertake headed for recognition progression, $e_u \in E$ represents to know as state its preparation framework, $\ker nal$ indicates its core role, ∂_u influence with the information, also λ designates the expanse starting the overexcited hydroplane, with $biclafun$ indicates binary categorization utility

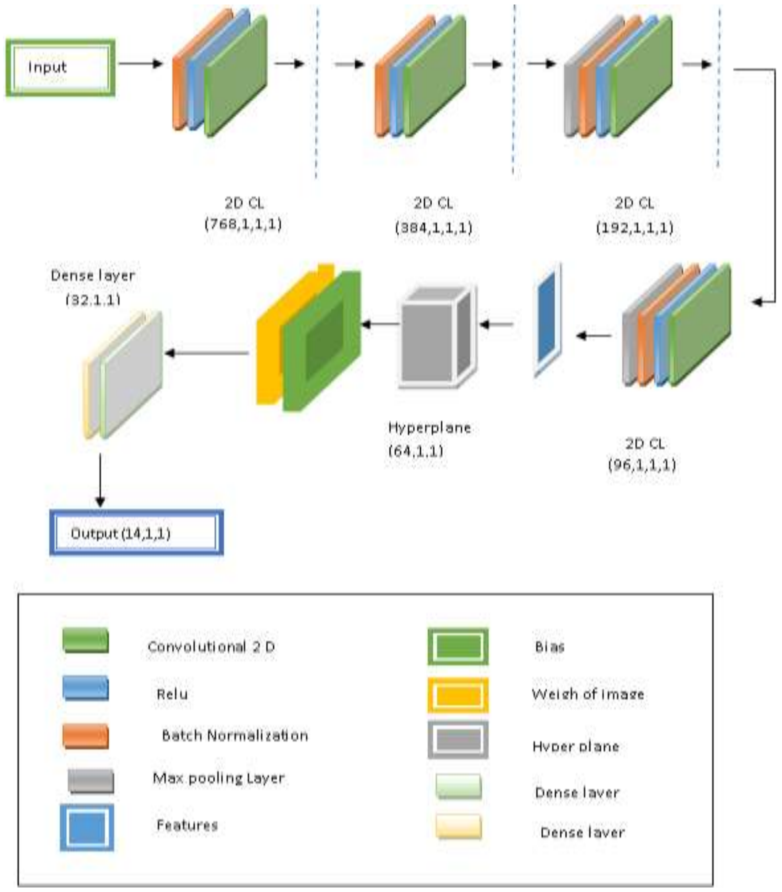

Fig.2 Architecture diagram of MB deep CNN

$$\min_{\partial} \frac{1}{2} \sum_{a=1}^{0} \sum_{b=1}^{0} \partial_a \partial_b \ker nal(e_a, e_b), \quad (3)$$

These multifaceted inputs are represented by a and b. A over formula is evaluated by

$$0 \leq \partial_a \leq \frac{1}{\beta n^{th}}, \sum_{a=1}^{0} \partial_a = 1, \quad (4)$$

Where $\beta \in [0,1]$ is used to express the normal parameter, the satisfied samples E_a is

$$\lambda = \sum_{b=1}^{0} \partial_b \ker nal(e_b, e_a), \quad (5)$$

λ learns from its information input investigation. Information classification in the deep stratum is dependent on the mathematical model's previously specified weights.

B. k-Fold Cross-Validation: Evaluating model performance is essential in machine learning to ensure practical effectiveness in real-world applications. Among various evaluation techniques, k-fold cross-validation has emerged as a particularly reliable and widely used approach. This method involves partitioning the dataset into *k* equally sized subsets, or folds, and then iteratively training the model on *k-1* folds while using the remaining fold for validation. This process is repeated *k* times, with each fold serving as the validation set exactly once, and the final performance metrics are derived by averaging the results across all iterations. Through using that complete information during simultaneous learning as examination, the k-fold cross-validation method minimises sample scepticism presenting a major benefit compared to standard train-test splits. It also aids in preventing overfitting by giving a more reliable indicator of a model's generalisability after many confirmation runs. Correctness, ability to recall, F1-score, and AUC-ROC are some of the most employing assessment criteria. Various metrics provide different observations based on the applications. Precise measures need required to design k-fold cross-validation, including information churning to avoid order-induced scepticism graded division to address disorganised information sets, including continuous monitoring achievement indicators throughout passes. To handle possible constraints consisting of overestimation in inadequate or complicated historical data, it is feasible to choose a suitable amount for *k*, usually among 5 and 10, which balances mathematical effectiveness alongside assessments precision. Additional strategies, among them interrelated a cross-valid or regularisation, can be used to address this problem. When it comes to human schooling, k-fold cross-validation provides a solid foundation for evaluating network effectiveness. This, in turn, allows for better judgements.

IV. RESULTS

Following is some example depicting the MB deep CNN classifier's effectiveness with a comparison of its assessment in graphical identification of events.

A. Experimental psychology, Configuration: Microsoft Windows 11 Graphics having 16 GB RAM in 1TB are used to execute the study.

B. Dataset Details: The collection of data that is used for academic purposes is: The study collection includes 4,90,000 photos, all of which depict different elements regarding human interactions. This information is matched.

C. Performance Metrics: Effectiveness indicators including empathy, precision, reliability, and accuracy are used to determine the model's performance. The deep CNN makes general right predictions based on accuracy measurements. With accuracy, we can determine how good the genuine prediction rate is, and with sensitivity, we can evaluate and check the positive examples produced by MB-Deep CNN. Measuring the actual negative rate is the specificity.

D. Experimental Outcome: Fig. 3 shows the picture-based output, which is shown in the research results subsection. The occurrence is identified using the features that were obtained, that include magnitude, orientation, as well as transition patterns in the photographic display.

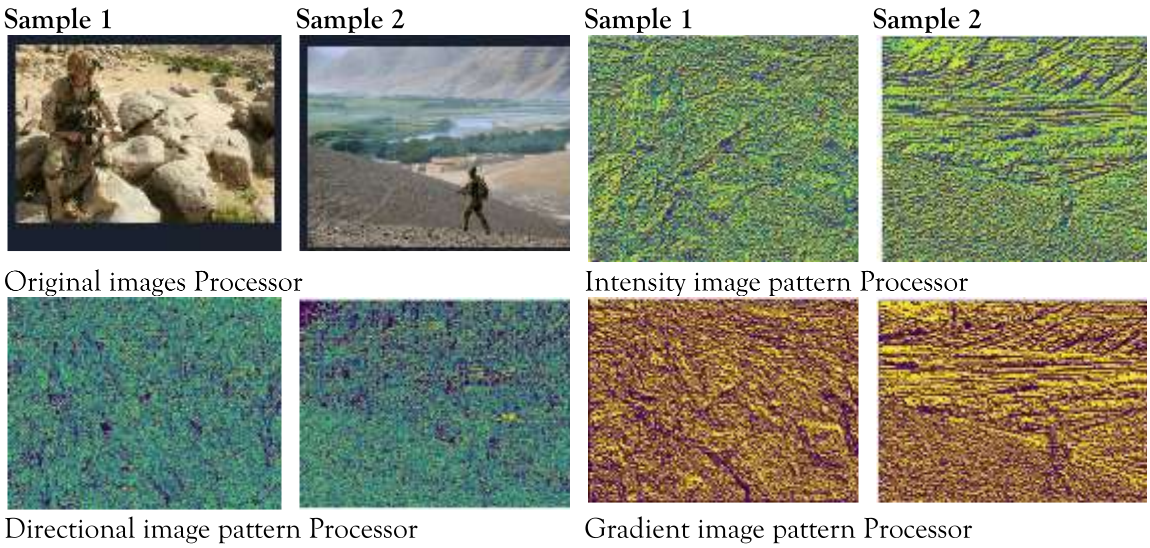

Fig.3 Various Image Pattern processors

E. Performance evaluation with k-fold: For k-fold 10 image-based event detection and reorganisation, the performance assessment of the deep CNN model is shown in Fig. 4. For epochs 100, 200, 300, 400, and 500, the MB-deep CNN model is accurate with k-fold 10 of 87.34%, 89.64%, 90.37%, 92.74%, and 95.72%, respectively. Over the course of the five epochs, the MB deep CNN model attains an accuracy of 88.81%, 89.42%, 91.42%, 91.63%, and 96.41%, measured in k-folds. Depending on the era, the sensitivity achieved by the MB deep CNN model with k-fold 10 is 92.06%, 92.32%, 92.84%, 94.78%, and 95.00%. Specifically, the MB deep CNN model attains k-fold 10 85.70%, 88.21%, 89.57%, 93.65%, and 96.44%.

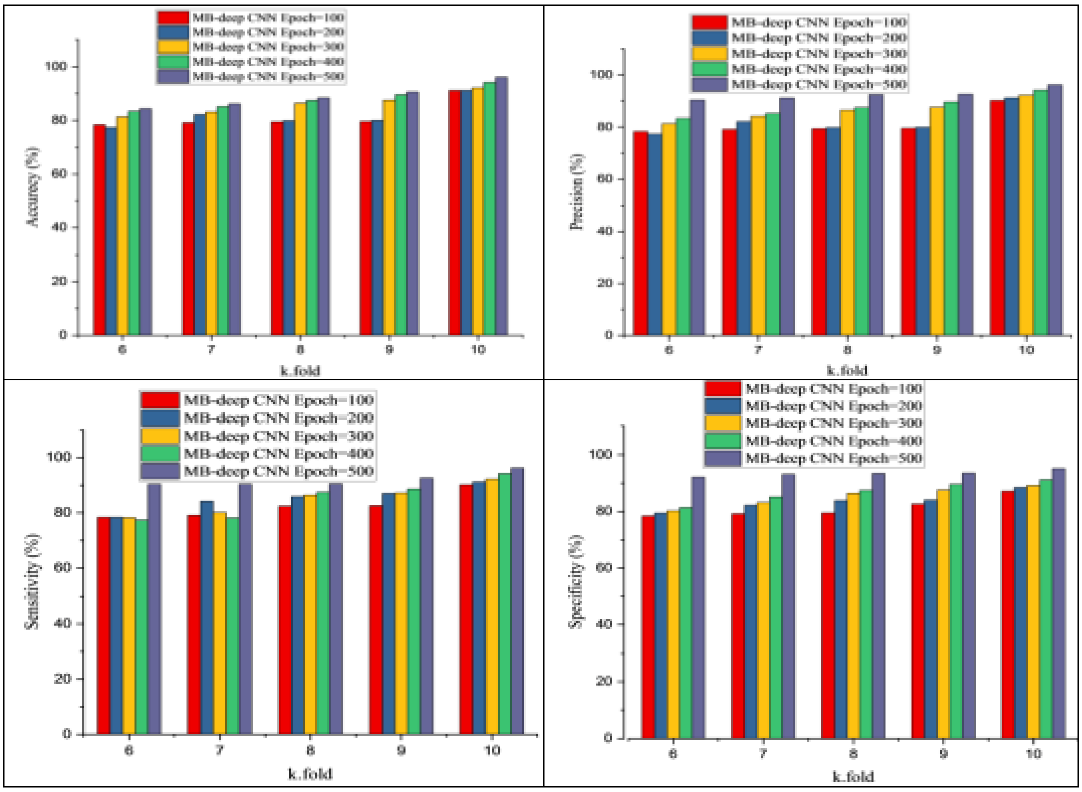

Fig. 4. Performance analysis with k-fold a) Accuracy, b) Precision, c) Sensitivity and d) Specificity

F. Comparative Analysis: Current methods such as DT [6], Multilayer perception (MLP) [35], HMM [8], 1Dimensional deep Convolutional Neural Network (1D deep CNN) [33], and 2Dimensional analysis deep Convolutional Neural Network (2D deep CNN) [34] are compared with the MB deep CNN model in this technique.Figure 6 shows how the MB deep CNN model compares to the conventional methods for k-fold 10. The performance measurements show that the MB-hyper plane model, with the help of the MB deep CNN with k-fold 10, increased the success rate. Compared to DT (11.09%), MLP (6.53%), HMM (3.32%), 1D deep CNN (3.82%), and 2D deep CNN (3.88%), the MB deep CNN model achieves an accuracy of 95.72%. Compared to the DT (15.70%), MLP (16.12%), HMM (18.68%), 1Dimensional deep convolutional neural network (10.83%), and 2Dimensional deep convolutional neural network (16.37%), the MB deep CNN achieved an accuracy of 96.41%. With an increase of 8.29% over DT, 12.30% over perception, 4.10% over HMM, 6.4% over I D deep CNN, and 4.17% over 2D deep CNN, the MB deep CNN model achieves a sensitivity of 95.00%. Compared to the DT (13.86%), MLP (0.86%), HMM (2.54%), 1D deep CNN (1.17%), and 2D deep CNN (13.49%), the MB deep CNN achieved a specificity of 96.44%.

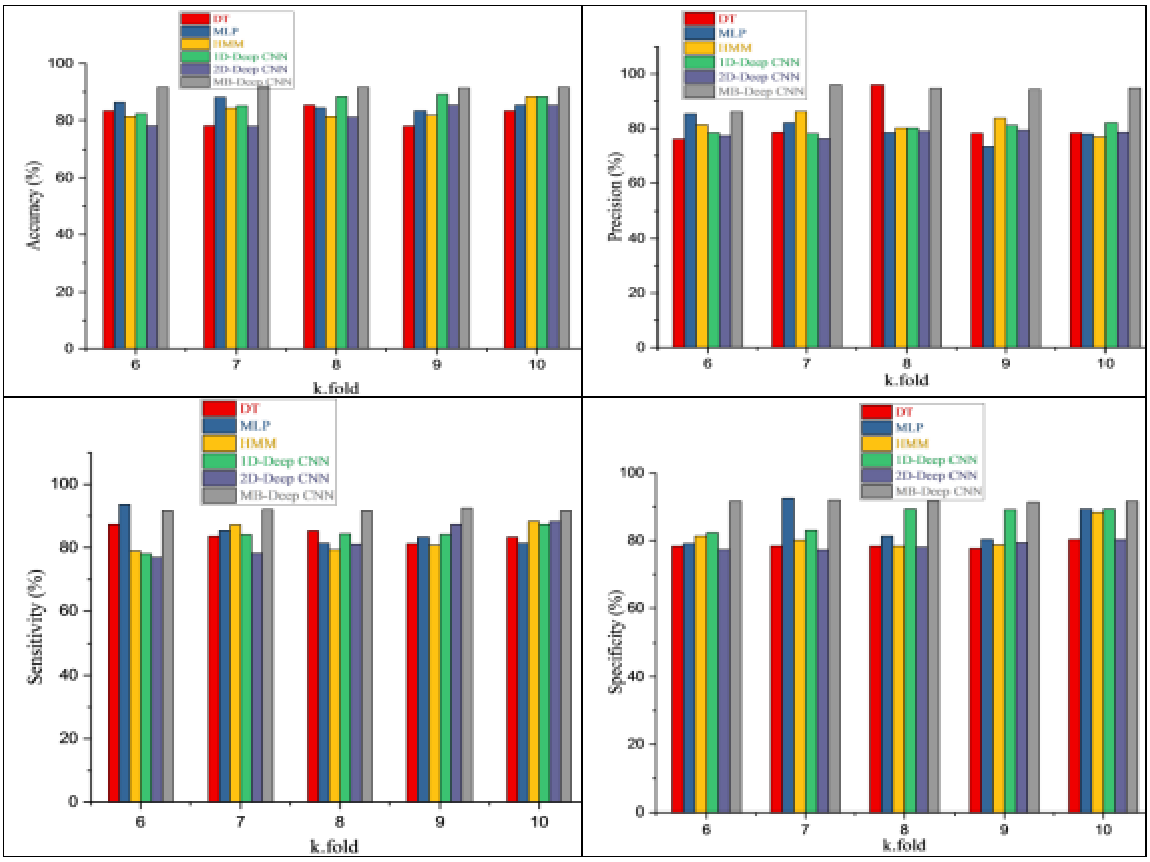

Fig. 5. Comparative analysis with k-fold a) Accuracy, b) Precision, c) Sensitivity and d) Specificity

G. *Comparative discussion:* In order to compare and contrast the MB deep CNN, which is used for event identification based on images, see Table 1. The DT approach was used for event detection, however it increased the data size, further complicating matters. We utilised the HMM method to describe what happened in the picture, but we didn't use it to detect every flaw. The event detection MLP needed data initialisation since it failed to correctly discriminate the event. The approach relies on 1D deep CNNs for event detection, but, the software used for this purpose was not optimal for detection. The event was detected using a 2D deep CNN, however due to its poor layer capacity, it was unable to address all difficulties. This study's only purpose is to identify and categorise events; nevertheless, the model gets over these problems by making use of the lowest boundary hyper plane layer that the SVM generates, hence the method's success rate is rather high.

Table 1: Comparative discussion

Methods		DT	MLP	HMM	1D deep CNN	2D deep CNN	MB deep CNN
Comparative discussion with K-fold	Precision (%)	83.18	95.12	93.89	95.09	83.08	96.15
	Sensitivity (%)	83.45	85.16	85.23	89.34	92.08	94.18
	Specificity (%)	83.18	95.16	93.89	95.23	83.14	96.25
	Accuracy (%)	83.29	85.26	92.24	89.56	92.02	95.24

V. Conclusion

The most cutting-edge deep Convolution Neural Network (CNN) model for event detection and recognition in pictures was introduced in this research. The use of 2-dimensional convolutions improves the model's capacity to represent the current attributes related to the previous or previous event detection characteristics. A hyperplane layer is constructed to divide the pictures using the Support Vector Machine (SVM). A distinct class is represented by each hyper plane layer split according to the bias and minimum/maximum weight. Dense layers help aggregate information from earlier convolution layers and bring the segmented pictures together. Accurate event detection is achieved by further partitioning the output of the convolution layers by the dense layer. The size and location of the event are determined by the characteristics of IIP, DIP, and GIP. This feature makes the event detection technique much more accurate. Precision(96.45%), sensitivity(94.68%), specificity (96.45%), and accuracy(94.68%) are all areas in which the MB deep CNN model surpasses the established method. To get the most out of the categorization model in the future, it's suggested to include bio-inspired algorithms and learning methodologies.

ACKNOWLEDGEMENT

All authors have accepted responsibility for the entire content of this manuscript and consented to its submission to the journal, reviewed all the results and approved the final version of the manuscript. S. Sai Satyanarayana Reddy designed the experiments and Seelam Meghana Reddy carried them out. Seelam Sanjana developed the model for image pereion . A. Seshappa prepared the manuscript with contributions from all co-authors.

REFERENCES

[1] Chong,, Yong Shean, and Yong Haur Tay. "Abnormal event detection in videos using spatiotemporal auto encoder." In Advances in Neural Networks-ISNN 2017: 14th International Symposium, ISNN 2017, Sapporo, Hakodate, and Muroran, Hokkaido, Japan, Proceedings, Part II 14, pp. 189-196. Springer International Publishing, June 2017.

[2] Medioni, Gérard, Isaac Cohen, François Brémond, Somboon Hongeng, and Ramakant Nevatia. "Event detection and analysis from video streams." IEEE Transactions on pattern analysis and machine intelligence 23, no. 8 873-889, 2001.

[3] Baillie, Mark, and Joemon M. Jose. "An audio-based sports video segmentation and event detection algorithm." In 2004 Conference on Computer Vision and Pattern Recognition Workshop, pp. 110-110. IEEE, 2004.

[4] Ding, Lei, Quanfu Fan, Jen-Hao Hsiao, and Sharath Pankanti. "Graph based event detection from realistic videos using weak feature correspondence." IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1262-1265. IEEE, 2010.

[5] Hongeng, Somboon, Ram Nevatia, and Francois Bremond. "Video-based event recognition: activity representation and probabilistic recognition methods." Computer Vision and Image Understanding 96, no. 2 129-162, 2004.

[6] Shyu, Mei-Ling, Zongxing Xie, Min Chen, and Shu-Ching Chen. "Video semantic event/concept detection using a subspace-based multimedia data mining framework." IEEE transactions on multimedia 10, no. 2252-259, 2008.

[7] Smeureanu, Sorina, Radu Tudor Ionescu, Marius Popescu, and Bogdan Alexe. "Deep appearance features for abnormal behavior detection in video." In Image Analysis and Processing-ICIAP 2017: 19th International Conference, Catania, Italy, Proceedings, Part II 19, pp. 779-789. Springer International Publishing, September 2017.

[8] Abou-Elailah, Abdal bassir, Valerie Gouet-Brunet, and Isabelle Bloch. "Detection of abrupt changes in spatial relationships in video sequences." In Pattern Recognition: Applications and Methods: 4th International Conference, ICPRAM 2015, Lisbon, Portugal, Revised Selected Papers 4, pp. 89-106. Springer International Publishing, January 2015.

[9] Ravanbakhsh, Mahdyar, Enverv Sangineto, Moin Nabi, and NicuSebe. "Training adversarial discriminators for cross-channel abnormal event detection in crowds." IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1896-1904. IEEE, 2019.

[10] J. S. Boreczky and L. D. Wilcox. A hidden Markov model framework for video segmentation using audio and image features. In Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP'98), pages 3741–3744, Seattle, WA, May 1998.

[11] [11] M. Baillie, J. M. Jose, and C. J. van Rijsbergen. Hmm model selection issues for soccer video. In Proc. 3rd International Conference of Image and Video Retrieval (CIVR'04), Dublin, EIRE, July 2004.

[12] Y. Wang, Z. Liu, and J. Huang. Multimedia content analysis using both audio and visual clues. In IEEE Signal Processing Magazine. IEEE, 2000.

[13] L Xie, S-F Chang, A Divakaran, and H Sun. Structure analysis of soccer video with hidden Markov models. In Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02). IEEE, 2002.

[14] V.Kettnaker, "Time-dependent HMMs for visual intrusion detection", Proc. IEEE Workshop on Detection and Recognizing Eventsin Video, 2003.

[15] H. Mousavi, S. Mohammadi, A. Perina, R. Chellali, and V. Murino. Analyzing for the detection of abnormal crowd behavior. In WACV, 2015.

[16] H.Rabiee, J. Haddadnia, H. Mousavi, M. Nabi, V. Murino, and N. Sebe. Emotion-based crowd representation for abnormality detection. ArXiv preprint arXiv:1607.07646, 2016.

[17] M. Sabokrou, M. Fayyaz, M. Fathy, and R. Klette. Deep cascade: Cascading 3d deep neural networks for fast anomaly detection and localization in crowded scenes. TIP,2017.

[18] V. Saligrama and Z. Chen. Video anomaly detection basedon local statistical aggregates. In CVPR, 2012.

[19] T. Salimans, I. J. Good fellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen. Improved techniques for training GANs. In NIPS, 2016.

[20] [N. Sebe, V. Murino, M. Ravanbakhsh, H. Rabiee,H. Mousavi, and M. Nabi. Abnormal event recognition in crowd environments. In Applied Cloud Deep Semantic Recognition. 2018.

[21] A. van den Oord and B. Schrauwen. Factoring variations in natural images with deep Gaussian Mixture Models. NIPS,2014.

[22] D. Xu, Y. Yan, E. Ricci, and N. Sebe. Detecting anomalous events in videos by learning deep representations of appearance and motion. CVIU, 2016.
[23] Zhou, Wei, Ling Zhang, Shengyu Gao, and Xin Lou. "Gradient-based feature extraction from raw bayer pattern images." IEEE Transactions on Image Processing 30 5122-5137, 2021.
[24] Vidal, Maider, and José Manuel Amigo. "Pre-processing of hyper spectral images. Essential steps before image analysis." Chemo metrics and Intelligent Laboratory Systems 117 138-148, 2012.
[25] Ghose, Shuvozit, Abhirup Das, Ayan Kumar Bhunia, and Partha Pratim Roy. "Fractional local neighborhood intensity pattern for image retrieval using genetic algorithm." Multimedia Tools and Applications 79, no. 25 18527-18552, 2020.
[26] Rivera, Adin Ramirez, Jorge Rojas Castillo, and Oksam Chae. "Local directional texture pattern image descriptor." Pattern Recognition Letters 51 pp94-100,2015.
[27] Othman, Abdullah, Naveed Iqbal, Sherif M. Hanafy, and Umair Bin Waheed. "Automated event detection and demising method for passive seismic data using residual deep convolutional neural networks." IEEE Transactions on Geosciences and Remote Sensing 60 pp1-11,2021.
[28] Attallah, Bilal, Amina Serir, and Youssef Chahir. "Feature extraction in palmprint recognition using spiral of moment skewers and kurtosis algorithm." Pattern Analysis and Applications 22, pp 1197-1205, 2019
[29] Cakir, Emre, Ezgi Can Ozan, and Tuomas Virtanen. "Filterbank learning for deep neural network based polyphonic sound event detection." In 2016 International Joint Conference on Neural Networks (IJCNN), pp. 3399-3406. IEEE, 2016.
[30] Kamarudin, Noraziahtulhidayu, S. A. R. Al-Haddad, Shaiful Jahari Hashim, and Mohammad Ali Nematollahi. "Feature extraction using spectral centroid and Mel frequency costrel coefficient for Quran accent automatic identification." In 2014 IEEE Student Conference on Research and Development, pp. 1-6. IEEE, 2014.
[31] Hu, Wei, Yangyu Huang, Li Wei, Fan Zhang, and Hengchao Li. "Deep convolutional neural networks for hyperspectral image classification." Journal of Sensors 2015 : 1-12,2015.
[32] Ghamisi, Pedram, Yushi Chen, and Xiao Xiang Zhu. "A self-improving convolution neural network for the classification of hyper spectral data." IEEE Geosciences and Remote Sensing Letters 13, no. 10, pp 1537-1541, 2016.
[33] Wu, Peng, Jing Liu, and Fang Shen. "A deep one-class neural network for anomalous event detection in complex scenes." IEEE transactions on neural networks and learning systems 31, no. 7 pp2609-2622, 2019.
[34] Escottá, Álvaro Teixeira, Wesley Beccaro, and Miguel Arjona Ramírez. "Evaluation of 1D and 2D deep convolutional neural networks for driving event recognition." Sensors 22, no. 11 4226,2022.
[35] Alam, Tasmia Ishrat, Imtiaz Mamun, Iftakhar Ali Khandokar, and Zubair Ahmed Anas. "Event detection and violence recognition from textual news through multilayer MLP and supervised learning." PhD diss., 2018.
[36] USER dataset "https://qualinet.github.io/databases/image/used a large scale social event detection dataset/" accessed on January, 2024.