

**SPECIMEN FORMAT FOR THESES OF MONTH**

<b>Faculty</b>	:	School of Engineering
<b>Department</b>	:	Computer Science and Engineering
<b>Branch/ Area:</b>	:	Deep Learning
<b>Sub Subject Heading:</b>	:	Video Anomaly Detection
<b>Candidate's Name</b>	:	Sreedevi R. Krishnan
<b>Candidate's Address with email</b>	:	Ammancheri House, Edavanakkad P.O. Ernakulam District, Kerala- 682502 19pheop005@avinuty.ac.in, sreedevirkrishnan@gmail.com
<b>Title of the thesis</b>	:	Hybrid Transfer Learning Models for Video Anomaly Detection in Surveillance Systems
(i) In Roman Script	-	-
(ii) In roman Script	-	-
<b>Nomenclature of Degree:</b>	:	Doctor Of Philosophy in Computer Science and Engineering
<b>Month &amp; Year of Enrolment:</b>	:	January 2020
<b>Month &amp; Year of Registration:</b>	:	January 2020
<b>Month &amp;Year of Submission:</b>	:	March 2025
<b>Month &amp;Year of Award</b>	:	July 2025
<b>Name of Supervisor</b>	:	Dr. P. Amudha
<b>Designation of Supervisor</b>	:	Professor
<b>Centre/department/school in which research was conducted</b>	:	Department of Computer Science and Engineering, School of Engineering
<b>University's Name &amp; Address</b>	:	Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu -641043

## **Abstract within 300 words:**

In a modern civilized community, public safety is of prime importance, and the detection of anomalous events has become a vital factor for a successful security system. Conventional Video Surveillance (VS) methods are inadequate due to the inability to analyze large sequential data. The growing demand for intelligent VAD motivates the need of automated anomaly detection.

The study develops and evaluates four hybrid models for VAD using Deep Learning techniques based on transfer learning. The first research phase proposed a hybrid CNN-YOLO model. This model uses CNN for model training and modified YOLOv4 for object detection, ensuring accurate and high-speed anomaly detection. This model processes a single random frame out of 100 input frames and yields a faster response.

To incorporate continuous video analysis, a hybrid model comprised of Residual Network (ResNet) and Long Short-Term Memory (LSTM) was executed in the second phase. ResNet-50 is employed for spatial feature extraction and LSTM to capture temporal relationships of the input video data even though this hybrid model enhances detection capability but has low accuracy, efficiency and generalization skill due to overfitting.

In the third phase, a segmentation-based anomaly detection model is implemented, by hybridizing Improved UNet (IUNet) with Cascade Sliding Window Technique (CSWT) reducing overfitting. The ConvLSTM replaces standard Convolution in IUNet for spatiotemporal feature extraction. CSWT estimates the anomaly score of the input video for classification. This model balances generalization and precision, achieving high accuracy with low false positives in crowded environment.

The fourth phase implemented a Hierarchical Multiscale-CNN with LSTM model, enhancing multi-scale feature identification and temporal data analysis. This model can work efficiently even in low-resolution video utilizing a Bilateral-Wave Denoise Technique. Multi-scale CNN is augmented by a Spatial Pyramid Pooling (SPP), which enhances the feature extraction. The performance of this model outperforms all the other models.

**i) Major objectives:**

- To design a model that utilizes feature extraction and object detection for faster VAD.
- To integrate spatial feature extraction and temporal sequence learning for continuous video analysis.
- To improve VAD by focusing on a spatiotemporal segmentation technique to improve accuracy and diminish overfitting.
- To enhance VAD in low-resolution video by capturing spatiotemporal dependencies and minimizing noise.

**ii) Hypothesis:**

The proposed hybrid models CNN-YOLO, ResNet-LSTM, IUNet-CSWT, Hierarchical M-CNN with LSTM and Spatio-Guard YOLO significantly enhance video anomaly detection by improving detection accuracy, spatio-temporal learning, anomaly localization and robustness against noise and low resolution thereby outperforming conventional methods in both frame-base and real-time surveillance scenarios.

**iii) Methodology:**

Apply hybrid Transfer Learning Techniques like CNN, YOLO, ResNet50, LSTM, UNet and Hierarchical Multiscale CNN for spatial and temporal feature extraction and classifies into normal and anomalous event.

**iv) Findings:**

- The CNN-YOLO hybrid model enables object detection with minimal hardware requirements, achieving an accuracy of 99.46%; however, the model is limited by processing only 100 frames, lacking continuous video analysis.
- The ResNet-LSTM model effectively captures spatio-temporal patterns in continuous video, ensuring consistent anomaly recognition, though it still suffers from overfitting issues.

- The IUNet-CSWT model achieved robust segmentation and enhanced PSNR of 35.04 dB through ConvLSTM and Wiener filtering, enabling accurate anomaly detection in noisy videos, though with a slight drop performance in low resolution video.
- Hierarchical M-CNN with LSTM demonstrated strong generalization and adaptability in low-resolution video with multiscale feature fusion and SPP, yielding 99.35% accuracy, 42.18dB PSNR, and 99.85% AUC, making it a balanced and high-performing model for intelligent surveillance.
- The Spatio-Guard YOLO model also demonstrates strong generalization and good real-time performance, marked by a 99.81% accuracy, validating its effectiveness for real-time anomaly detection.

## **Examiners**

### **Internal Examiner:**

Dr B Santhosh Kumar, Professor & Head,  
Department of CSE  
Guru Nanak Institute of Technology,  
Ibrahimpattanam,  
Ranga Reddy (Dist),  
Telangana- 501506

### **External Examiner:**

Dr. Subarna Shakya,  
Professor Electronics and Computer Engineering,  
Pulchowk Campus,  
Institute of Engineering,  
Tribhuvan University,  
Nepal. 44700