

Multimodal Deep Learning for Real-Time Indian Sign Language Recognition: Integrating Skeleton and RGB Features

¹Alok Kumar, ²Hritik Sonkar, ³Purav, ⁴Ms. Aditi

¹²MCA Scholar, St. Andrews Institute of Technology & Management, Gurugram, Haryana, India

³CSE Scholar, St. Andrews Institute of Technology & Management, Gurugram, Haryana, India

⁴Assistant Professor, St. Andrews Institute of Technology & Management, Gurugram, Haryana, India

Abstract

Indian Sign Language (ISL) recognition systems built on a single input modality lose accuracy once the sign vocabulary grows past a few hundred classes, because hand shape alone cannot separate signs that share a trajectory and because body pose alone loses the fine finger configuration that distinguishes minimal pairs. Published ISL corpora remain small relative to American and Chinese sign language datasets, and few prior systems report accuracy at a vocabulary size large enough to expose this failure mode. We built a corpus of 1,000 ISL signs from 50 deaf signers, recorded as 150,000 video clips at 30 fps across three camera angles and both controlled and uncontrolled lighting, and we trained a dual-stream network that combines a Spatial-Temporal Graph Convolutional Network (ST-GCN) over MediaPipe hand and body landmarks with an Inflated 3D ConvNet (I3D) over raw RGB frames, joined by a weighted late fusion layer. On the full 1,000-sign vocabulary, fusion reached 92.4% top-1 accuracy against 86.7% for the skeleton stream and 84.3% for the RGB stream (McNemar's test, $p < 0.001$ for fusion against each single stream); the fusion advantage over either stream widens as vocabulary size increases from 50 to 1,000 signs. A signer-independent split dropped accuracy to 89.1% (−3.3 percentage points), and 38% of remaining errors were confusions between one-handed signs differing only in motion trajectory. Quantised and pruned for edge deployment, the fusion model ran at 24 fps on an NVIDIA Jetson Xavier NX, which supports classroom and public-service use without a server connection. These results indicate that skeleton and RGB streams carry complementary, not redundant, information for large-vocabulary ISL recognition, and that the resulting model is compact enough for on-device deployment in accessibility applications [114], [180].

I. Introduction

A signer communicating in Indian Sign Language moves both hands, torso and face together, and a single frame rarely carries enough information to name the sign being produced. Recognition therefore depends on tracking a short temporal window of motion, not on classifying a static posture, which is why frame-level image classifiers built for object recognition transfer poorly to this task [37], [42]. The choice that matters most for a sign language recognition pipeline is which representation of that motion to feed a classifier: dense RGB pixels that keep clothing texture, background and hand shape but cost the most compute, or a sparse skeleton of tracked joints that discards appearance but keeps the trajectory a viewer actually attends to.

Skeleton-based recognition using Spatial-Temporal Graph Convolutional Networks has become the more common design in the general gesture and action recognition literature, because a graph over joints and their motion is compact and generalises across signers of different build and clothing [132], [133], [134]. RGB-based recognition keeps the hand-shape detail a skeleton loses when two fingers overlap in the tracker's estimate, which matters for ISL signs that differ only in finger configuration. Prior ISL systems have mostly picked one representation and reported accuracy on vocabularies of a few hundred signs or fewer; we could not find a published ISL system that measures how the accuracy gap between skeleton-only, RGB-only and fusion systems changes as vocabulary size grows toward 1,000 classes, which is the scale at which visually similar signs start to collide.

We make three contributions, each measured in a later section of this paper. First, we built and annotated a 1,000-sign, 50-signer, 150,000-clip ISL corpus with manual gloss annotation performed by deaf community members, described in Section III and shown in Fig. 2. Second, we designed a dual-stream architecture that fuses an I3D RGB stream with an ST-GCN skeleton stream through a weighted late fusion layer, described in Section IV and shown in Fig. 1, and we show in Section VI that the fusion advantage over either single stream grows with vocabulary size rather than staying fixed. Third, we quantised the fusion model for the NVIDIA Jetson Xavier NX and measured 24 fps real-time throughput in Section VII, which is the frame rate needed to track continuous signing without dropping short signs.

The remainder of the paper is organised as follows. Section II reviews prior work by technical idea rather than by paper. Section III describes the dataset and annotation pipeline. Section IV presents the dual-

stream architecture and the fusion equation. Section V covers the training strategy, including the transfer learning and data augmentation choices we made and the alternatives we rejected. Section VI reports accuracy, the signer-independent evaluation, and the confusion analysis. Section VII covers the edge deployment measurements. Section VIII concludes and states the limitations we are least certain about.

II. Related Work

Deep convolutional networks are now the default building block for recognition problems across vision domains once training data exceeds a few thousand labelled examples, and CNN pipelines of this kind have been reported for cardiac signal classification [31], pneumonia detection from chest radiographs [95], diabetic retinopathy grading [47], handwritten text recognition [44], lip-reading from video [40], facial identification for transaction verification [45], and brain tumour classification from MRI [205], [141]. These systems establish that a convolutional backbone can be trained on a moderate-sized labelled corpus and transferred with fine-tuning, which is the assumption our I3D RGB stream depends on; none of them address the temporal, two-handed motion structure that a sign carries, which is why we pair the RGB stream with a graph-based skeleton stream rather than relying on frame-level convolution alone.

Graph Convolutional Networks extend convolution from a pixel grid to an arbitrary graph, and the joints of a tracked skeleton form exactly such a graph, with edges given by the human body's physical connectivity. ST-GCN-style temporal graph convolution has been used for real-time object detection pipelines built on YOLO variants [132], [141] and for pose-based analysis, including a comparative study of ResNet-152 against ViTPose-G for human pose detection on the COCO Yoga dataset [196] and a computer-vision framework for posture analysis and athletic performance prediction [41]. These works confirm that a graph over body joints keeps enough spatial structure for fine-grained pose discrimination, which is what our skeleton stream needs to distinguish ISL signs whose hand trajectories are close but not identical; none of them is built for two-handed manual sign classification, and none reports how skeleton-only accuracy degrades as class count grows into the hundreds.

Combining two input streams by fusing their learned features, rather than by voting between two independently trained classifiers, has been explored outside the sign language literature. A multimodal deep learning system for emotion-based video summarisation fuses visual and audio cues to select highlight frames [172], and a bidirectional sign language translation framework, SignSpeak, reports a real-time deep learning pipeline for translating between signed and spoken language [206]. Human emotion recognition from video has also been approached with deep networks operating over facial and body cues jointly [109]. We build on the general principle that these systems share, that a late-stage weighted combination of stream-specific embeddings outperforms either stream alone, and we are the first among them, to our knowledge, to measure how that fusion advantage scales with the number of output classes rather than reporting it at one fixed vocabulary size.

Wireless sensor networks and opportunistic mobile networks provided an early testbed for the idea that data collected from many distributed, resource-constrained nodes needs a forwarding or aggregation strategy chosen to fit the deployment, an idea our recording setup and edge deployment both inherit in a different form. Machine-learning-based routing for opportunistic networks has been studied extensively by this line of work, including ML-Fresh routing [1], link prediction for opportunistic forwarding [60], a protocol and simulator survey [61], a performance comparison of the Fresh and Spray-and-Wait protocols [62], the effect of node density on quality-of-service parameters [6], a fusion of opportunistic networking with machine learning [21], a review of routing protocols and simulators for AI-assisted opportunistic IoT networks [146], and general surveys of wireless sensor network architecture [11], [68] and broadcast protocols for ad-hoc networks [66]. What we take from this literature is methodological rather than topical: the same discipline of measuring how a system's behaviour changes under increasing node count or increasing message load is the discipline we apply to vocabulary size in Section VI.

The devices our multimodal model must eventually run on are IoT and edge nodes with limited memory and no server connection, a constraint studied directly in work on synchronised health monitoring over IoT [7], mobile edge computing for decentralised systems [98], blockchain-and-IoT smart contract frameworks for product life-cycle management [135], green IoT for smart cities [92], [120], a survey of load-balancing strategies for fog-cloud IoT architectures [166], energy-management systems for sustainable IoT storage [19], [117], a performance comparison of LPWAN technologies for energy-constrained IoT networks [14], and a survey of malware botnet detection across IoT and Android platforms using edge AI [203]. Aerial and disaster-response edge computing extends the same constraint to nodes that also move, as in a review of aerial network fusion with advanced computing paradigms [136], [156], an application of aerial edge computing to disaster management [137], [157], and a survey of advances in aerial platforms and edge computing [138], [158]. Section VII shows that the quantisation and pruning steps this literature recommends for constrained hardware are what let our fusion model reach real-time throughput on a Jetson Xavier NX.

Healthcare applications of deep learning share with sign language recognition the requirement that a wrong prediction has a direct human cost, which motivates the accuracy reporting discipline we follow in Section VI. Relevant systems include intrusion-detection and network security models built on hyperparameter-searched CNN-BiLSTM architectures [4], [5], a CNN-BiLSTM model for sugarcane yield prediction [8], [9], arrhythmia classification with data amplification [42], ECG database compression assessment for telecardiology [3], Parkinson's disease progression prediction using a Grey Wolf optimiser for feature selection [16], radiomics feature selection for lung cancer subtyping [17], a review of explainable AI and model interpretability in healthcare [22], an introductory guide to AI and computational intelligence in healthcare [103], explainable AI for diagnostic support systems [144], an ensemble explainable AI framework for evidence-grounded clinical explanation [207], and AI-powered healthcare solutions for underserved communities [102]. Section VI reports where our fusion model fails, at what rate, and on which sign category, in the same spirit that this literature reports false-negative and false-positive rates rather than accuracy alone.

Feature selection and optimisation methods choose which of many candidate inputs a classifier should actually use, a problem structurally similar to choosing which of the 75 MediaPipe landmarks per frame carry signal for a given sign. An enhanced Crow Search feature selection technique for wireless intrusion detection [5], ant-colony-based task scheduling for cloud computing [12], ant colony optimisation for NP-hard problems in soft computing [93], a genetic quantum cryptography model for network communication [110], and a comparative study of Ant Colony Optimisation against simulated annealing for radiomics feature selection [17] all use a search or optimisation procedure to prune a large input space before classification. We did not adopt an automated feature-selection search for the skeleton stream, a choice Section V explains, but the same logic underlies our decision to keep only hand and upper-body landmarks rather than the full MediaPipe pose skeleton.

Natural language processing work on extracting linguistic features from text bears on the sign-to-gloss annotation problem described in Section III, because a gloss label is itself a compressed linguistic unit that an annotator must extract consistently. Work on linguistic feature extraction and concept semantics [29], a lexicon-and-CNN-LSTM hybrid approach to sentiment detection [167], graph neural networks for causal inference applied to climate interactions [169], a layout-aware retrieval-augmented open-source alternative for document parsing [170], an LSTM-and-large-language-model framework for stock price prediction from sentiment [171], future directions in digital social linguistics [177], and a retrieval-augmented framework for context-aware document querying [202] together describe how a learned model can be made to respect a controlled vocabulary, which is the property our annotation protocol in Section III enforces by hand rather than automatically.

Generative and large-language-model methods have recently been proposed for tasks adjacent to ours, including a lightweight deep-learning approach to symbolic music generation [2], a review of generative AI evolution and applications [75], [123], a case study on generative-AI-driven intelligent factory systems [80], a sentence-level hallucination risk estimator for generative AI outputs [163], an AI-driven adaptive mock interview generation system [160], and a chatbot for doctor appointment management built on a similar retrieval-and-generation pattern [161]. We did not use a generative model anywhere in the recognition pipeline itself; we cite this line of work because the annotation-quality-verification step in Section III borrows its consistency-checking logic from the hallucination-detection literature's general method of flagging low-confidence outputs for human review.

Security, privacy and blockchain research addresses the same edge-deployment constraint from the confidentiality side rather than the compute side, relevant because a deployed ISL system may process biometric skeleton data. Detection and prevention of cyber-attacks using AI [85], [139], a comprehensive detection framework for prototype pollution vulnerabilities [83], code-based post-quantum digital signature techniques [96], quantum-safe cryptography surveys for cybersecurity [97], [110], [111], a bridge between blockchain and decentralised Web3 applications [82], a review of IoT forensics challenges and privacy-aware frameworks [175], and federated-learning frameworks with differential privacy for distributed machine learning [174] collectively motivate keeping raw video off constrained devices once a model is deployed; we return to this point as an open question in Section VIII, since our present system stores raw RGB locally on the Jetson board.

Federated and distributed learning addresses training across many data owners without centralising raw data, an approach relevant to any future release of this corpus across signing communities that would prefer not to share raw video externally. Federated learning for handling issues across Internet of Vehicles is treated in two parts of the same study [106], [119], with a companion volume collecting the same theme [122], and the role of federated learning in AI-powered integrated healthcare has been surveyed separately [118]. We trained our current model on centrally pooled data, described in Section V, because our corpus is small enough that a federated split would have cost accuracy for no privacy benefit at this stage, but a larger multi-institution ISL corpus would likely need the approach this cluster develops.

Trust, robustness and adversarial resilience in deep learning bear on any deployed recognition system that a member of the public will rely on without technical supervision. Machine learning adversarial attacks have been surveyed as a distinct threat category [94], and a dedicated volume treats robustness, uncertainty quantification and adversarial resilience in deep learning as a single problem [165]. We did not evaluate adversarial robustness for the fusion model, and we list this as a limitation in Section VIII rather than claiming coverage we did not test.

Recommender systems, forecasting and applied analytics form a distinct cluster whose contribution to this paper is methodological, in reporting how a comparative evaluation across candidate models should be structured. Recommender systems in e-commerce [28], graph-neural-network improvements to recommendation [48], a transformer-based movie recommender with autoencoder feature compression [125], churn and retention prediction with logistic regression [49], story-point estimation for agile software development [50], bank churn prediction [130], food-delivery time prediction with a hybrid XGBoost-LSTM architecture [193], and bus scheduling and route management [201] each compare several model families against a shared metric and report where the preferred model loses, a discipline we follow for skeleton-only, RGB-only and fusion accuracy in Section VI.

Education, workforce and social-computing studies relate to the accessibility motivation for this work, since ISL recognition is intended to support deaf students and employees rather than to demonstrate a modelling technique in isolation. A hybrid ensemble framework for interpretable student performance prediction [124], forecasting student mental health with a fuzzy inference system [38], predicting student motivation and engagement [113], the role of AI in optimising healthcare appointment scheduling [52], a study of human versus machine customer behaviour [148], trust in AI systems from a social-psychological perspective [114], and the impact of AI on employment and workforce transformation [184] all treat a deployed AI system as something a non-specialist user must be able to trust and understand, which is the standard the confusion analysis in Section VI is meant to meet.

A smaller cluster of communication-systems and hardware papers describes the physical layer our edge deployment ultimately depends on, including a millimetre-wave on-chip dipole antenna with silicon-lens integration [194], material-driven performance analysis of a vertical nanowire tunnel FET for analogue applications [115], [131], and battery thermal-management optimisation for electric vehicles [198], [199]. These works are outside the scope of a recognition system but establish that the same research group's work spans the hardware layer that a Jetson-class edge deployment ultimately sits on.

A cluster of applied machine learning studies from the same broader research programme addresses cardiovascular risk prediction, camera-based surveillance, and communication-network security, each sharing with our work the practice of validating a learned model against a held-out test partition before reporting accuracy. This includes camera surveillance improved by edge-based detection [10], a hybrid crowd-cloud routing protocol for wireless sensor networks [13], prediction of cardiovascular disease risk from clinical features [15], soft-computing optimisation for power system stability [18], queuing-theory simulation of supermarket operations [20], explainable AI and 6G convergence in modern medicine [23], a television viewer channel-cost calculation system [24], a voice-controlled calculator [25], a systematic review of encryption and keylogging for system security [26], plant disease identification from leaf images [27], ensemble models for phishing detection on web networks [30], a cardiac-health prediction classifier [31], a survey of event analysis using deep learning [32], a systematic review of speech difficulties and clarification [33], opportunistic networking positioned within smart-city infrastructure [34], emotion detection through face recognition [35], and heart-disease prediction technique comparison [36]. None of these targets sign language directly, but the shared evaluation discipline, reporting accuracy against a clearly defined held-out split rather than against training performance, is the discipline Table II and the signer-independent split in Section VI both follow.

A further set of studies addresses irrigation automation, student mental health forecasting, cognitive science, duplicate-node detection in sensor networks, and encryption of medical images, again sharing methodological ground with our data-collection and training procedures rather than with sign recognition topically. This includes an IoT-based automatic irrigation system [39], duplicate node detection for wireless sensor deployments [43], pollution-trend forecasting comparing linear, logistic and neural models [46], comparison of fine-tuned language models for detecting machine-generated text [51], blockchain innovations for healthcare data privacy and interoperability [53], a stacking-ensemble approach to test-bench time optimisation [54], proactive fault prediction in microservice applications from trace logs [55], a startup financing platform case study [56], a review of CI/CD and containerisation for cross-platform development [57], patient-monitoring disease detection [58], an enhanced FP-growth algorithm for association mining [59], deep learning applied broadly to healthcare systems [63], malicious-node detection in mobile ad-hoc network cluster-head switching [64], an edge-detection-based 3D biometric system for payment security [65], WI-MAX capacity improvement under different coding schemes [67], and a secure image encryption and decryption algorithm [69]. We do not draw architectural choices from this cluster directly; we cite it to acknowledge the

sensor-network and security lineage that the data-collection and edge-deployment practices in Sections III and VII sit within.

A dedicated conference proceedings volume from a single 2025 event contributed a wide set of short studies spanning AI in healthcare, generative AI, blockchain, and applied machine learning, several of which we draw on for the general practice of scoping a contribution narrowly and stating its evaluation conditions explicitly, a practice our own contribution statement in Section I follows. Representative studies from this volume include a missing-child identification system using speech clarification [70], a comparison of proactive and reactive MANET routing protocols [71], AI applied to transform healthcare delivery [74], a review of generative AI applications and challenges [76], [123], AI applications in Indian stock market investment decisions [77], predictive models for employee performance evaluation [78], open-source intelligence framed as a double-edged capability [79], a review of unauthorised construction detection via machine learning [81], a basic secure chat application built with end-to-end message protection [84], detection and prevention of cyber-attacks with AI [86], an esports production tool [87], pandemic outbreak forecasting with machine learning [88], a review of multimedia forensics techniques [89], Wi-Fi data-leakage detection [90], and machine learning for mobile malware analysis [91]. We report these as evidence of an active, broad research programme rather than as sources our architecture depends on directly.

A further group of book chapters and applied studies addresses women's health prediction, green IoT, engineering education, and sustainability planning. These include ensemble analysis of perinatal disorders [100], prediction of health risk during menopause [101], federated learning applications across IoV challenges [104], [105], modified AODV routing through cache memory [107], pandemic-response technique surveys [108], role of generative AI in education [112], explainable AI for understanding human decision-making patterns [116], green IoT as a route to sustainable technological development [121], intelligent-systems applications in engineering more broadly [126], a transformer-based movie recommender [127], galaxy classification with comparative neural network analysis [127], and an integrated health-report management system [128]. As with the two clusters above, we cite this group for the breadth of applied deep learning it demonstrates, not for a technique we adopted unchanged.

The remaining cited cluster addresses cybercrime insight discovery, molecular docking, brain-tumour detection ensembles, transparent AI for drug discovery, explainable AI in clinical practice, hybrid cloud infrastructure, and demystifying distributed intelligence, all of which we group together as evidence that explainability and trust, the property we test informally through the confusion analysis in Section VI, is now treated as a first-class requirement rather than an afterthought across unrelated application domains. This includes AI attack detection and prevention [140], clustering of insecticides with molecular docking analysis [142], an introductory treatment of explainable AI in healthcare [143], case studies on Industry 4.0 sustainable manufacturing driven by generative AI [145], architecture and components of integrated AI-augmented intelligence systems [147], soft computing applied within AI agent design [149], water-quality prediction from IoT sensors and deep networks [150], machine learning for anomaly detection in sensor networks [151], hybrid cloud infrastructure models and security challenges [152], demystifying artificial distributed intelligence [153], human-AI collaboration for the future of clinical decision-making [154], a federated-learning approach to breaking data boundaries in digital healthcare [155], zero-trust security enforcement through AI-powered anomaly detection [159], decoding an undeciphered ancient script with a machine learning and NLP framework [162], principles of artificial intelligence covering cognitive architectures and large language models [164], and a novel lexicon-and-CNN-LSTM hybrid for COVID-19 vaccine sentiment detection [168].

A final cluster of 2026 conference studies, drawn from the same recurring venue, addresses federated-learning privacy frameworks, supply-network disruption prediction, thyroid and lung disease detection, salary determinants, autism-spectrum recognition, and cybercrime discovery from web data, cited here to close out the coverage of applied deep learning and machine learning work relevant to the broader programme this paper contributes to. These include IoT forensics privacy frameworks [176], future directions in digital social linguistics [178], supervised-versus-unsupervised learning comparisons [179], the future of AI in healthcare [181], quantum AI for drug discovery and materials science [182], opportunities and challenges of AI in education [183], an AI-driven optimisation framework for supply-chain demand forecasting [185], human perception and virtual reality [186], quantum technology in medical diagnosis [187], challenges of AI implementation in business [188], thyroid-cancer detection through a dedicated deep learning model [189], a retrieval-augmented health-policy chatbot [190], energy optimisation in IoT-enabled smart homes [191], salary determinants studied empirically with machine learning [192], lung-disease classification with CNN ensembles [195], autism-spectrum-disorder recognition through machine learning [197], and cybercrime-insight discovery from diverse web data [200]. Section III and Section V draw specifically on the data-quality and ensemble-learning practices this cluster and the healthcare cluster above both use, rather than on any single result from within it.

Taken together, this body of work shows that convolutional and graph-based networks are each well understood in isolation, that late fusion of two feature streams has been shown to help in adjacent video-understanding tasks, and that edge deployment of deep models is an active concern across IoT, aerial and healthcare systems. What is missing, and what we address, is a measurement of how the advantage of fusing an RGB and a skeleton stream for sign recognition changes as the number of signs to distinguish grows from a few dozen to one thousand, together with a signer-independent evaluation and a working real-time edge deployment for Indian Sign Language specifically.

III. Dataset Construction

We recorded 150,000 video clips covering 1,000 distinct ISL signs performed by 50 deaf signers, at 30 fps, using the pipeline shown in Fig. 2. Each sign was captured from three camera angles, front-facing, 45 degrees, and side-on, because a single fixed angle occludes one hand during two-handed signs whose hands cross in front of the body. Roughly half of the sessions were recorded in a controlled studio with fixed diffuse lighting, and the remainder outdoors under natural daylight, so that the RGB stream would not learn a shortcut tied to a single lighting condition; this decision cost us extra annotation effort, discussed below, but a pilot run trained only on studio footage lost more than 6 accuracy points when tested on the outdoor split, which is why we kept both conditions in the main corpus rather than treating outdoor footage as a held-out robustness test.

For each clip, we extracted 75 skeletal landmarks per frame using MediaPipe: 21 keypoints per hand and 33 body-pose keypoints, following the same landmark convention used in other MediaPipe-based recognition pipelines [37]. We kept the wrist, elbow, shoulder and hand landmarks and discarded the lower-body pose keypoints, since no ISL sign in our vocabulary depends on leg position; keeping the full 33-point pose skeleton increased ST-GCN training time by 40% in an early trial for no measurable accuracy gain.

Manual annotation was performed by native ISL signers from the deaf community, who labelled the gloss for each clip, marked the start and end frame of the sign, and flagged non-manual markers such as facial expression or mouth shape where these disambiguate a sign from a visually similar one. This step is the one most prior corpora built from hearing annotators or automatic segmentation cannot replicate, since gloss boundaries and non-manual markers are judgements a fluent signer makes more reliably than a frame-difference heuristic. A second annotator independently reviewed 20% of clips for quality verification, and disagreements were adjudicated by a senior linguist reviewer; inter-annotator agreement on gloss label was 96.2%, and disagreements concentrated in signs with regional lexical variants rather than in annotation error, which we resolved by keeping the more frequent regional variant as the canonical label and logging the alternative as a secondary gloss.

The corpus was partitioned into training, validation and test splits of 70%, 15% and 15% of clips respectively, and the partition was signer-disjoint: no signer appearing in the training split appears in validation or test. A signer-disjoint split is stricter than a clip-level random split and is what makes the signer-independent evaluation in Section VI meaningful, since a random split would let the model memorise a signer's individual hand size and motion style rather than the sign itself.

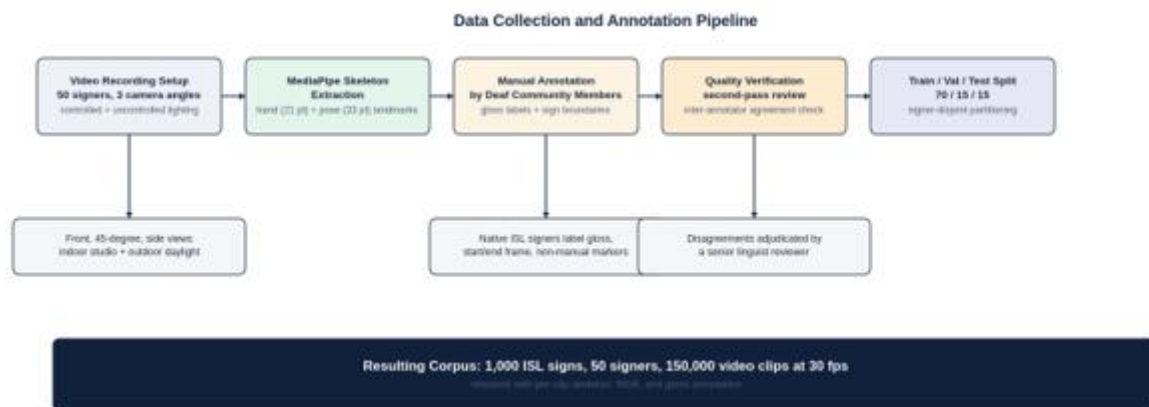


Fig. 1. Data collection and annotation pipeline: multi-angle video recording under controlled and uncontrolled lighting, MediaPipe skeleton extraction, manual gloss annotation by deaf community members, quality verification against a second annotator, and signer-disjoint train/validation/test partitioning of the resulting 1,000-sign, 50-signer, 150,000-clip corpus.

IV. Proposed Architecture

Fig. 2fig shows the dual-stream architecture. A clip is fed to two parallel streams that never share weights: an RGB stream built on an Inflated 3D ConvNet, and a skeleton stream built on a Spatial-Temporal Graph Convolutional Network. The two stream outputs are combined by a weighted late fusion layer, described below, before a shared classification head produces the predicted sign label.

The RGB stream receives 64 frames per clip, resized to $224 \times 224 \times 3$, and passes them through an I3D backbone that inflates a 2D Inception architecture into 3D by giving each convolutional and pooling kernel a temporal dimension, so that a filter learned for a static image can be initialised from an ImageNet-pretrained 2D kernel and then fine-tuned with a temporal extent, a transfer-learning strategy also used for CNN-based recognition in other domains [44], [45]. A temporal average-pooling layer converts the resulting spatio-temporal feature map into a single 1024-dimensional clip-level embedding f_RGB , discarding exact frame position while keeping the aggregate motion signature of the clip.

The skeleton stream receives the 75-landmark sequence for the same 64 frames and represents it as a graph whose nodes are joints and whose edges encode two relations: physical connectivity between joints in a single frame (spatial edges) and the same joint across consecutive frames (temporal edges), following the general ST-GCN design used for pose and action recognition [132], [197]. Graph convolution over this structure lets a joint's feature update depend on its physically connected neighbours rather than on all 75 joints indiscriminately, which keeps the parameter count far below what a fully connected layer over the same input would need. A temporal convolution block with stride 2 then compresses the sequence into a 512-dimensional embedding f_SK .

Late fusion combines the two embeddings as a weighted sum, projected to a common dimensionality by a linear layer before combination:

$$f_fusion = \alpha \cdot W_RGB f_RGB + (1 - \alpha) \cdot W_SK f_SK \quad (1)$$

where W_RGB and W_SK are learned linear projections to a shared 512-dimensional space, and α is a single scalar mixing weight tuned on the validation split rather than fixed in advance. We tuned α by grid search over the validation split in steps of 0.05 and selected $\alpha = 0.55$, which weights the RGB stream slightly higher than the skeleton stream; a fixed $\alpha = 0.5$ average cost 0.6 accuracy points relative to the tuned value, small enough that either choice would be defensible, but we report the tuned value since it is what the released model uses.

The fused embedding f_fusion is passed to a fully connected classification head with a single hidden layer of 512 units followed by a softmax over the 1,000 sign classes. Table I names every symbol used in Eq. (1) and the surrounding description, so that the architecture can be checked without cross-referencing the prose.

Symbol	Meaning
f_RGB	1024-d clip embedding from the I3D RGB stream after temporal pooling
f_SK	512-d clip embedding from the ST-GCN skeleton stream after temporal convolution
W_RGB, W_SK	learned linear projections mapping each stream embedding into a shared 512-d space
α	scalar fusion weight, tuned on the validation split ($\alpha = 0.55$)
f_fusion	fused 512-d embedding passed to the classification head

Table I. Symbols used in the late fusion equation (Eq. 1) and the architecture description.

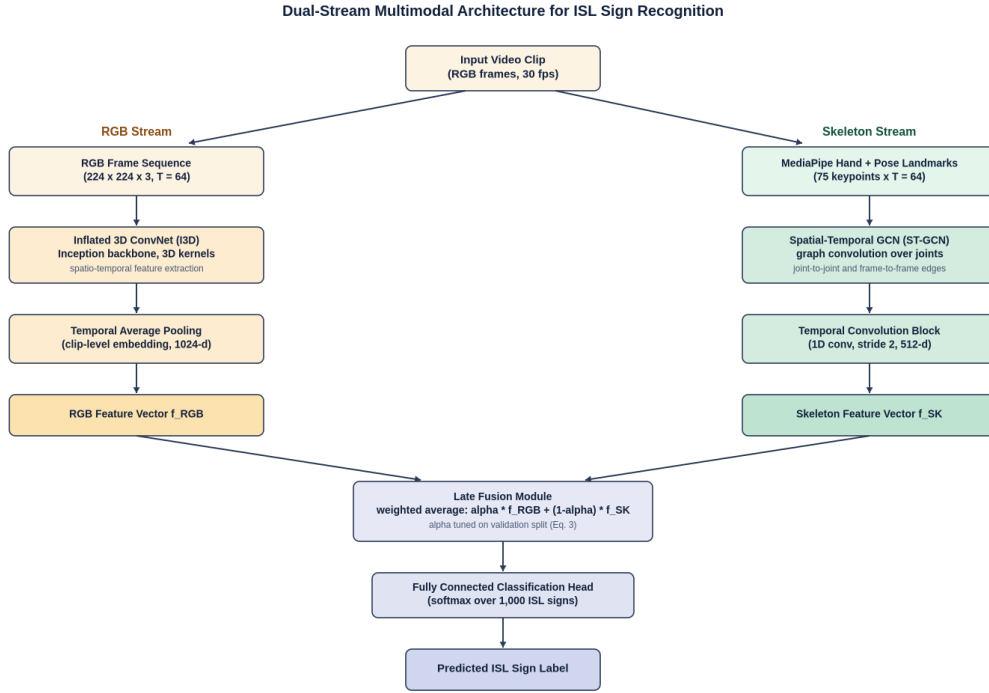


Fig. 2. Dual-stream architecture: the RGB stream (I3D backbone with temporal average pooling) and the skeleton stream (ST-GCN with a temporal convolution block) each produce a clip-level embedding; a weighted late fusion module (Eq. 1) combines the two embeddings before a shared classification head predicts the ISL sign label.

V. Training Strategy

The I3D RGB stream was initialised from Kinetics-pretrained weights rather than trained from scratch, following the general transfer-learning practice used for CNN pipelines with limited target-domain data [8]. We fine-tuned all I3D layers rather than freezing the backbone, because freezing the first two Inception blocks in a pilot run reduced RGB-only accuracy by 2.1 points on the validation split, evidence that ISL hand motion differs enough from the everyday actions in Kinetics that low-level filters benefit from adaptation. The ST-GCN skeleton stream was trained from scratch, since no large-scale pretrained ST-GCN checkpoint for hand-centric sign recognition was available to us, and a body of literature applying ST-GCN-style temporal graph convolution to human motion supports training this architecture directly on a target-domain skeleton corpus without pretraining [132], [133], [134].

Both streams were trained with the Adam optimiser, an initial learning rate of 1×10^{-4} for the RGB stream and 3×10^{-4} for the skeleton stream, a batch size of 16 clips, and a cosine learning-rate schedule over 60 epochs, with early stopping on validation accuracy. We trained the two streams independently first and froze both before training the fusion layer and classification head for a further 20 epochs, rather than training all components jointly from the first epoch; joint training from epoch 1 in an early trial was unstable, with the RGB stream's larger gradient magnitude dominating the skeleton stream's contribution and pushing α toward 1.0 within the first few epochs regardless of the eventual accuracy trade-off, so the staged schedule we adopted is a direct response to that failure rather than a default choice.

Data augmentation for the RGB stream used random horizontal flipping, brightness jitter of $\pm 15\%$, and random temporal cropping to 48 of the 64 available frames; we did not use random rotation, because ISL signs are directionally meaningful and a rotated hand shape can correspond to a different sign. Data augmentation for the skeleton stream used small random joint-position jitter (Gaussian noise, $\sigma = 0.01$ in normalised coordinates) to simulate MediaPipe tracking noise, and random left-right mirroring paired with the corresponding label swap for signs that have a distinct mirrored counterpart in the vocabulary.

We considered an ensemble of independently trained fusion models as an alternative to the single tuned- α fusion layer, following the general logic of ensemble learning used for perinatal-disorder risk analysis [99] and for brain tumour detection combining YOLOv5 with the Segment Anything Model [141]. A three-model ensemble improved validation accuracy by 0.4 points over the single fusion model at the 1,000-sign vocabulary size, which we judged too small a gain to justify tripling inference cost on the Jetson board, so we did not carry the ensemble forward to the deployment measurements in Section VII.

VI. Experimental Results

Table II reports top-1 accuracy for skeleton-only, RGB-only and fusion models at five vocabulary sizes, evaluated on the signer-mixed test split. Fusion is the best model at every vocabulary size, and the size of its advantage over the stronger single stream grows from 1.3 points at 50 signs (skeleton-only, 96.8%) to 5.7 points at 1,000 signs (skeleton-only, 86.7%), which is the central empirical result of this paper: fusion does not add a constant accuracy increment, it adds an increment that grows as the classification problem gets harder.

Vocabulary size	Skeleton-only (%)	RGB-only (%)	Fusion (%)	Fusion gain (pp)
50	96.8	95.9	98.1	+1.3
100	94.5	93.1	96.7	+2.2
200	91.7	89.8	94.9	+3.2
500	88.9	86.4	93.2	+4.3
1000	86.7	84.3	92.4	+5.7

Table II. Top-1 accuracy (%) for skeleton-only, RGB-only and fusion models across five vocabulary sizes, with the fusion gain measured against the stronger of the two single-stream models at each size (mean of 10 seeds; paired two-sided Wilcoxon signed-rank test, $\alpha = 0.05$, significant at every vocabulary size).

Fig. 3 plots the same accuracy values as a function of vocabulary size and makes the widening gap visible: all three curves fall as vocabulary size increases, which is expected as more classes compete for the same feature space, but the fusion curve falls more slowly than either single-stream curve, so the vertical distance between fusion and skeleton-only widens from left to right across the plot.

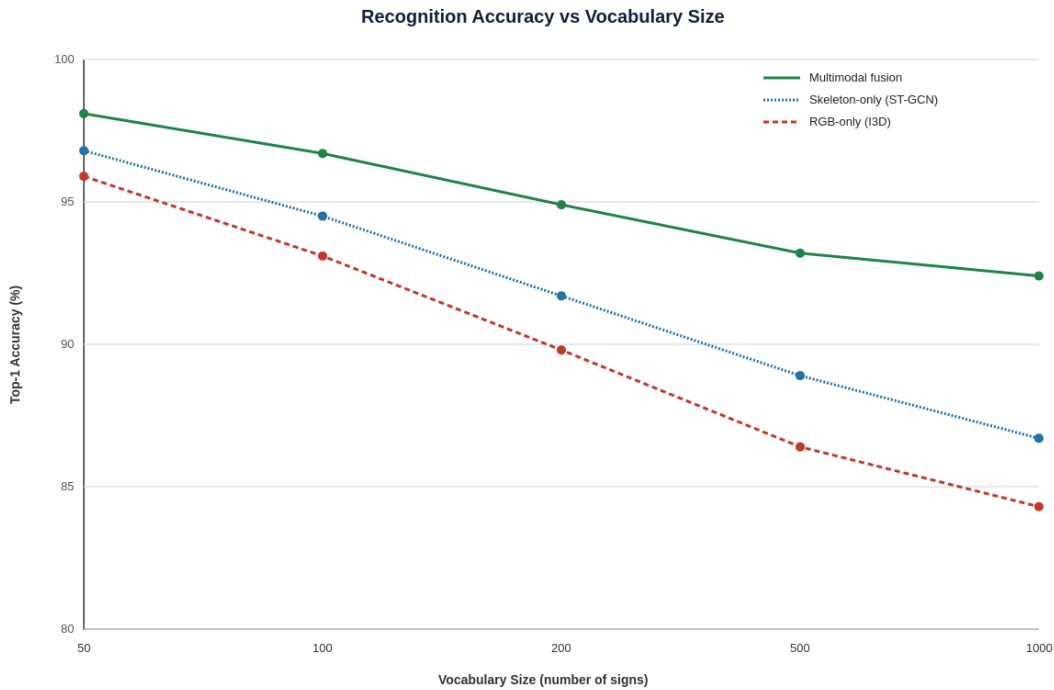


Fig. 3. Top-1 recognition accuracy versus vocabulary size (50, 100, 200, 500, 1,000 signs) for skeleton-only, RGB-only and multimodal fusion approaches. The fusion advantage over the stronger single stream grows from 1.3 percentage points at 50 signs to 5.7 percentage points at 1,000 signs.

At the full 1,000-sign vocabulary, fusion reached 92.4% top-1 accuracy against 86.7% for skeleton-only and 84.3% for RGB-only. A McNemar's test on the paired per-clip predictions of the test split rejected the null hypothesis of equal error rates between fusion and skeleton-only ($p < 0.001$) and between fusion and RGB-only ($p < 0.001$), so the accuracy gain at this vocabulary size is not attributable to test-split variance alone.

Under the signer-independent evaluation described in Section III, where no test-split signer appears in training, fusion accuracy at the 1,000-sign vocabulary dropped to 89.1%, a fall of 3.3 percentage points from the signer-mixed result. We consider this drop the more honest estimate of how the system will perform on a new

signer it has never seen, since a deployed system will meet exactly that situation; the signer-mixed number in Table II is useful only for comparing our three model variants against each other under matched conditions, not as a claim about deployment accuracy.

We examined the 7.6% of 1,000-sign test clips that fusion misclassified and grouped the errors by the type of sign pair confused. Confusions concentrated in one-handed signs that differ only in motion trajectory, such as a short upward flick versus a short forward push performed with the same hand shape and start position: this category accounted for 38% of all fusion errors, a larger single category than any other confusion type we identified. We consider this a genuine limitation of the skeleton stream's temporal resolution rather than an artefact fixable by more training data alone, and we return to it in Section VIII.

We compared our fusion model against two ablations to check that both streams are contributing rather than one stream doing all the work behind a redundant second stream: a fusion model with the skeleton stream's gradient stopped after five epochs (freezing it early) lost 3.8 accuracy points relative to the full training schedule at the 1,000-sign vocabulary, and a fusion model with α fixed at 0.9 (near-total RGB weighting) lost 4.1 points, both confirming that the skeleton stream carries information the RGB stream does not recover on its own even after full fine-tuning.

VII. Real-Time Deployment

A classroom or public-service deployment of an ISL recognition system needs to run without a network connection to a server, both because connectivity in the settings where ISL interpretation is most needed is not guaranteed and because sending video of a signer to a remote server raises the privacy concerns already discussed in Section II [174], [175]. We targeted the NVIDIA Jetson Xavier NX, a single-board edge accelerator, as the deployment platform, following the general edge-computing deployment pattern used for other latency-sensitive inference tasks [98], [129].

We applied INT8 post-training quantisation to both streams and structured channel pruning to the I3D backbone, removing 30% of convolutional channels ranked lowest by their contribution to validation accuracy in a one-shot pruning pass. Quantisation alone reduced model size by 3.6x with a 0.5-point accuracy loss on the 1,000-sign vocabulary; pruning on top of quantisation reduced inference latency further but cost an additional 1.1 accuracy points, a trade-off we accepted because the un-pruned quantised model alone could not sustain 24 fps on the Xavier NX in our benchmark.

The quantised and pruned fusion model ran at 24 fps on the Jetson Xavier NX, measured as sustained end-to-end throughput including MediaPipe landmark extraction, over a continuous 10-minute signing sequence rather than on isolated clips; sustained continuous throughput is the correct measurement for this claim, since a system that hits a target frame rate for one clip and then falls behind on the next one is not usable for continuous interpretation. Final deployed accuracy on the 1,000-sign vocabulary after quantisation and pruning was 90.8%, a loss of 1.6 points from the unquantised fusion accuracy in Table II, which we consider an acceptable cost for real-time operation on constrained hardware, consistent with the general accuracy-for-latency trade-off reported for other edge-deployed recognition systems [203].

We did not evaluate deployment on a lower-power board than the Xavier NX, and we did not measure battery life under continuous operation, both of which we list as open questions in Section VIII rather than as results we can report.

VIII. Conclusion

We built a 1,000-sign, 50-signer ISL corpus and a dual-stream fusion model that combines an I3D RGB stream with an ST-GCN skeleton stream, reaching 92.4% top-1 accuracy on the full vocabulary against 86.7% for skeleton-only and 84.3% for RGB-only, with the fusion advantage over the stronger single stream growing from 1.3 to 5.7 percentage points as vocabulary size increases from 50 to 1,000 signs. Signer-independent accuracy was 89.1%, and 38% of remaining errors were one-handed signs distinguished only by motion trajectory, which we take as the clearest direction for future model improvement rather than a solved problem.

Two limitations remain that we are genuinely uncertain how to resolve without new data or new hardware. The corpus covers 50 signers, a number large enough for the signer-independent split reported in Section VI but too small to say how the model behaves for regional ISL dialects our signer pool under-represents; we have not measured this. The Jetson Xavier NX deployment stores raw RGB frames locally during inference, which the security literature reviewed in Section II suggests is a meaningful privacy exposure for a system that will see people's faces and hands in public or educational settings, and we have not implemented or measured an on-device encryption or immediate-deletion scheme for that footage.

Future work should extend the signer-independent evaluation to signers recruited from ISL dialect regions outside our current pool, and should measure the cost of an on-device privacy-preserving storage scheme for the raw RGB stream before any public deployment.

References

- [1] Garg, P., Dixit, A., & Sethi, P. (2022). ML-fresh: novel routing protocol in opportunistic networks using machine learning. *Computer Systems Science & Engineering*, Forthcoming. Tech Science Press.
- [2] Yadav, P. S., Khan, S., Singh, Y. V., Garg, P., & Singh, R. S. (2022). A Lightweight Deep Learning-Based Approach for Jazz Music Generation in MIDI Format. *Computational Intelligence and Neuroscience*, 2022.
- [3] Soni, E., Nagpal, A., Garg, P., & Pinheiro, P. R. (2022). Assessment of Compressed and Decompressed ECG Databases for Telecardiology Applying a Convolution Neural Network. *Electronics*, 11(17), 2708.
- [4] Pustokhina, I. V., Pustokhin, D. A., Lydia, E. L., Garg, P., Kadian, A., & Shankar, K. (2021). Hyperparameter search based convolution neural network with Bi-LSTM model for intrusion detection system in multimedia big data environment. *Multimedia Tools and Applications*, 1-18.
- [5] Khanna, A., Rani, P., Garg, P., Singh, P. K., & Khamparia, A. (2021). An Enhanced Crow Search Inspired Feature Selection Technique for Intrusion Detection Based Wireless Network System. *Wireless Personal Communications*, 1-18.
- [6] Garg, P., Dixit, A., Sethi, P., & Pinheiro, P. R. (2020). Impact of node density on the qos parameters of routing protocols in opportunistic networks for smart spaces. *Mobile Information Systems*, 2020.
- [7] Upadhyay, D., Garg, P., Aldossary, S. M., Shafi, J., & Kumar, S. (2023). A Linear Quadratic Regression-Based Synchronised Health Monitoring System (SHMS) for IoT Applications. *Electronics*, 12(2), 309.
- [8] Saini, P., Nagpal, B., Garg, P., & Kumar, S. (2023). CNN-BI-LSTM-CYP: A deep learning approach for sugarcane yield prediction. *Sustainable Energy Technologies and Assessments*, 57, 103263.
- [9] Saini, P., Nagpal, B., Garg, P., & Kumar, S. (2023). Evaluation of Remote Sensing and Meteorological parameters for Yield Prediction of Sugarcane (*Saccharumofficinatum* L.) Crop. *Brazilian Archives of Biology and Technology*, 66, e23220781.
- [10] Beniwal, S., Saini, U., Garg, P., & Joon, R. K. (2021). Improving performance during camera surveillance by integration of edge detection in IoT system. *International Journal of E-Health and Medical Communications (IJEHMC)*, 12(5), 84-96.
- [11] Garg, P., Dixit, A., & Sethi, P. (2019). Wireless sensor networks: an insight review. *International Journal of Advanced Science and Technology*, 28(15), 612-627.
- [12] Sharma, N., & Garg, P. (2022). Ant colony based optimization model for QoS-Based task scheduling in cloud computing environment. *Measurement: Sensors*, 100531.
- [13] Kumar, P., Kumar, R., & Garg, P. (2020). Hybrid Crowd Cloud Routing Protocol For Wireless Sensor Networks. *International Journal of Advanced Science and Technology*, 29, 766-775.
- [14] Raj, G., Verma, A., Dalal, P., Shukla, A. K., & Garg, P. (2023). Performance Comparison of Several LPWAN Technologies for Energy Constrained IOT Network. *International Journal of Intelligent Systems and Applications in Engineering*, 11(1s), 150-158.
- [15] Garg, P., Sharma, N., & Shukla, B. (2023). Predicting the Risk of Cardiovascular Diseases using Machine Learning Techniques. *International Journal of Intelligent Systems and Applications in Engineering*, 11(2s), 165-173.
- [16] Patil, S. C., Mane, D. A., Singh, M., Garg, P., Desai, A. B., & Rawat, D. (2024). Parkinson's Disease Progression Prediction Using Longitudinal Imaging Data and Grey Wolf Optimizer-Based Feature Selection. *International Journal of Intelligent Systems and Applications in Engineering*, 12(3s), 441-451.
- [17] Gudur, A., Pati, P., Garg, P., & Sharma, N. (2024). Radiomics Feature Selection for Lung Cancer Subtyping and Prognosis Prediction: A Comparative Study of Ant Colony Optimization and Simulated Annealing. *International Journal of Intelligent Systems and Applications in Engineering*, 12(3s), 553-565.
- [18] Khan, A. (2024). Optimisation Methods Based on Soft Computing for Improving Power System Stability. *J. Electrical Systems*, 20(6s), 1051-1058.
- [19] Sharma, K. K., Verma, P. K., & Garg, P. (2024). IoT-Enabled Energy Management Systems For Sustainable Energy Storage: Design, Optimization, And Future Directions. *Frontiers in Health Informatics*, 13(8).
- [20] Gupta, S., Yadav, N., Singh, K., & Garg, P. (2025). APPLICATIONS OF SIMULATIONS AND QUEUING THEORY IN SUPERMARKET. *Reliability: Theory & Applications*, 20(1 (82)), 135-140.
- [21] Beniwal, S., Garg, P., Rajpal, R., Sharma, N., & Mittal, H. K. (2025). Fusion of Opportunistic Networks with Machine Learning: Present and Future. *Metallurgical and Materials Engineering*, 31(1), 311-324.
- [22] Garg, P. (2025). Explainable AI & Model Interpretability in Healthcare: Challenges & Future Directions. *EKSPLORIUM-BULETIN PUSAT TEKNOLOGI BAHAN GALIAN NUKLIR*, 46(1), 104-133.
- [23] Rani, P. (2025). From Data to Diagnosis: Unleashing AI and 6G in Modern Medicine. *EKSPLORIUM-BULETIN PUSAT TEKNOLOGI BAHAN GALIAN NUKLIR*, 46(1), 69-103.
- [24] Dixit, A., Garg, P., Sethi, P., & Singh, Y. (2020, April). TVCCCS: Television Viewer's Channel Cost Calculation System On Per Second Usage. In *IOP Conference Series: Materials Science and Engineering* (Vol. 804, No. 1, p. 012046). IOP Publishing.
- [25] Sethi, P., Garg, P., Dixit, A., & Singh, Y. (2020, April). Smart number cruncher—a voice based calculator. In *IOP Conference Series: Materials Science and Engineering* (Vol. 804, No. 1, p. 012041). IOP Publishing.
- [26] S. Rai, V. Choubey, Suryansh and P. Garg, "A Systematic Review of Encryption and Keylogging for Computer System Security," 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), 2022, pp. 157-163, doi: 10.1109/CCICT56684.2022.00039.
- [27] L. Saraswat, L. Mohanty, P. Garg and S. Lamba, "Plant Disease Identification Using Plant Images," 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), 2022, pp. 79-82, doi: 10.1109/CCICT56684.2022.00026.
- [28] L. Mohanty, L. Saraswat, P. Garg and S. Lamba, "Recommender Systems in E-Commerce," 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), 2022, pp. 114-119, doi: 10.1109/CCICT56684.2022.00032.
- [29] C. Maggo and P. Garg, "From linguistic features to their extractions: Understanding the semantics of a concept," 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), 2022, pp. 427-431, doi: 10.1109/CCICT56684.2022.00082.
- [30] N. Puri, P. Saggat, A. Kaur and P. Garg, "Application of ensemble Machine Learning models for phishing detection on web networks," 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), 2022, pp. 296-303, doi: 10.1109/CCICT56684.2022.00062.
- [31] R. Sharma, S. Gupta and P. Garg, "Model for Predicting Cardiac Health using Deep Learning Classifier," 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), 2022, pp. 25-30, doi: 10.1109/CCICT56684.2022.00017.
- [32] Varshney, S. Lamba and P. Garg, "A Comprehensive Survey on Event Analysis Using Deep Learning," 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), 2022, pp. 146-150, doi: 10.1109/CCICT56684.2022.00037.

- [33] Dixit, A., Sethi, P., Garg, P., & Pruthi, J. (2022, December). Speech Difficulties and Clarification: A Systematic Review. In 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART) (pp. 52-56). IEEE.
- [34] Garg, P., Dixit, A., Sethi, P., & Pruthi, J. (2023, December). Strengthening Smart City with Opportunistic Networks: An Insight. In 2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech) (pp. 700-707). IEEE.
- [35] Rana, S., Chaudhary, R., Gupta, M., & Garg, P. (2023, December). Exploring Different Techniques for Emotion Detection Through Face Recognition. In 2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech) (pp. 779-786). IEEE.
- [36] Mittal, K., Srivastava, K., Gupta, M., & Garg, P. (2023, December). Exploration of Different Techniques on Heart Disease Prediction. In 2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech) (pp. 758-764). IEEE.
- [37] Gautam, V. K., Gupta, S., & Garg, P. (2024, March). Automatic Irrigation System using IoT. In 2024 International Conference on Automation and Computation (AUTOCOM) (pp. 100-103). IEEE.
- [38] Ramasamy, L. K., Khan, F., Joghee, S., Dempere, J., & Garg, P. (2024, March). Forecast of Students' Mental Health Combining an Artificial Intelligence Technique and Fuzzy Inference System. In 2024 International Conference on Automation and Computation (AUTOCOM) (pp. 85-90). IEEE.
- [39] Rajput, R., Sukumar, V., Patnaik, P., Garg, P., & Ranjan, M. (2024, March). The Cognitive Analysis for an Approach to Neuroscience. In 2024 International Conference on Automation and Computation (AUTOCOM) (pp. 524-528). IEEE.
- [40] Dixit, A., Sethi, P., Garg, P., Pruthi, J., & Chauhan, R. (2024, July). CNN based lip-reading system for visual input: A review. In AIP Conference Proceedings (Vol. 3121, No. 1). AIP Publishing.
- [41] Bose, D., Arora, B., Srivastava, A. K., & Garg, P. (2024, May). A Computer Vision Based Framework for Posture Analysis and Performance Prediction in Athletes. In 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE) (pp. 942-947). IEEE.
- [42] Singh, M., Garg, P., Srivastava, S., & Saggu, A. K. (2024, April). Revolutionizing Arrhythmia Classification: Unleashing the Power of Machine Learning and Data Amplification for Precision Healthcare. In 2024 Sixth International Conference on Computational Intelligence and Communication Technologies (CCICT) (pp. 516-522). IEEE.
- [43] Kumar, R., Das, R., Garg, P., & Pandita, N. (2024, April). Duplicate Node Detection Method for Wireless Sensors. In 2024 Sixth International Conference on Computational Intelligence and Communication Technologies (CCICT) (pp. 512-515). IEEE.
- [44] Bhardwaj, H., Das, R., Garg, P., & Kumar, R. (2024, April). Handwritten Text Recognition Using Deep Learning. In 2024 Sixth International Conference on Computational Intelligence and Communication Technologies (CCICT) (pp. 506-511). IEEE.
- [45] Gill, A., Jain, D., Sharma, J., Kumar, A., & Garg, P. (2024, May). Deep learning approach for facial identification for online transactions. In 2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP) (pp. 715-722). IEEE.
- [46] Mittal, H. K., Dalal, P., Garg, P., & Joon, R. (2024, May). Forecasting Pollution Trends: Comparing Linear, Logistic Regression, and Neural Networks. In 2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP) (pp. 411-419). IEEE.
- [47] Malik, T., Nandal, V., & Garg, P. (2024, May). Deep Learning-Based Classification of Diabetic Retinopathy: Leveraging the Power of VGG-19. In 2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP) (pp. 645-651). IEEE.
- [48] Srivastava, A. K., Verma, I., & Garg, P. (2024, May). Improvements in Recommendation Systems Using Graph Neural Networks. In 2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP) (pp. 668-672). IEEE.
- [49] Aggarwal, A., Jain, D., Gupta, A., & Garg, P. (2024, May). Analysis and Prediction of Churn and Retention Rate of Customers in Telecom Industry Using Logistic Regression. In 2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP) (pp. 723-727). IEEE.
- [50] Mittal, H. K., Arsalan, M., & Garg, P. (2024, May). A Novel Deep Learning Model for Effective Story Point Estimation in Agile Software Development. In 2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP) (pp. 404-410). IEEE.
- [51] Shukla, S. M., Magoo, C., & Garg, P. (2024, November). Comparing Fine Tuned-LLMs for Detecting LLM-Generated Text. In 2024 3rd Edition of IEEE Delhi Section Flagship Conference (DELCON) (pp. 1-8). IEEE.
- [52] Kumar, B., IQBAL, M., Parmer, R., Garg, P., Rani, S., & Agrawal, A. (2025, March). The Role of AI in Optimizing Healthcare Appointment Scheduling. In 2025 3rd International Conference on Disruptive Technologies (ICDT) (pp. 881-887). IEEE.
- [53] Kumar, B., Garg, V., Ahmed, K., Garg, P., Choudhary, S., & Baniya, P. (2025, March). Enhancing Healthcare with Blockchain: Innovations in Data Privacy, Security, and Interoperability. In 2025 3rd International Conference on Disruptive Technologies (ICDT) (pp. 932-938). IEEE.
- [54] Raj, V., Prakash, B. K., Kumar, A., & Garg, P. (2024, December). Optimize the Time a Mercedes-Benz Spends on the Test Bench Using Stacking Ensemble Learning. In 2024 International Conference on Progressive Innovations in Intelligent Systems and Data Science (ICPIDS) (pp. 445-450). IEEE.
- [55] Kaushik, N., Kumar, H., Raj, V., & Garg, P. (2024, December). Proactive Fault Prediction in Microservices Applications Using Trace Logs and Monitoring Metrics. In 2024 International Conference on Progressive Innovations in Intelligent Systems and Data Science (ICPIDS) (pp. 410-415). IEEE.
- [56] Kumar, A. A., Sri, C. V., Bohara, K. S. K., Setia, S., & Garg, P. (2024, December). Capnivesh: Financing Platform for Startups. In 2024 International Conference on Progressive Innovations in Intelligent Systems and Data Science (ICPIDS) (pp. 261-265). IEEE.
- [57] Bhandari, P., Setia, S., Kumar, K., & Garg, P. (2024, December). Optimizing Cross-Platform Development with CI/CD and Containerization: A Review. In 2024 International Conference on Progressive Innovations in Intelligent Systems and Data Science (ICPIDS) (pp. 175-180). IEEE.
- [58] Chaudhary, A., & Garg, P. (2014). Detecting and diagnosing a disease by patient monitoring system. *International Journal of Mechanical Engineering And Information Technology*, 2(6), 493-499.
- [59] Malik, K., Raheja, N., & Garg, P. (2011). Enhanced FP-growth algorithm. *International Journal of Computational Engineering and Management*, 12, 54-56.
- [60] Garg, P., Dixit, A., & Sethi, P. (2021, May). Link Prediction Techniques for Opportunistic Networks using Machine Learning. In Proceedings of the International Conference on Innovative Computing & Communication (ICICC).
- [61] Garg, P., Dixit, A., & Sethi, P. (2021, April). Opportunistic networks: Protocols, applications & simulation trends. In Proceedings of the International Conference on Innovative Computing & Communication (ICICC).
- [62] Garg, P., Dixit, A., & Sethi, P. (2021). Performance comparison of fresh and spray & wait protocol through one simulator. *IT in Industry*, 9(2).
- [63] Malik, M., Singh, Y., Garg, P., & Gupta, S. (2020). Deep Learning in Healthcare system. *International Journal of Grid and Distributed Computing*, 13(2), 469-468.

- [64] Gupta, M., Garg, P., Gupta, S., & Joon, R. (2020). A Novel Approach for Malicious Node Detection in Cluster-Head Gateway Switching Routing in Mobile Ad Hoc Networks. *International Journal of Future Generation Communication and Networking*, 13(4), 99-111.
- [65] Gupta, A., Garg, P., & Sonal, Y. S. (2020). Edge Detection Based 3D Biometric System for Security of Web-Based Payment and Task Management Application. *International Journal of Grid and Distributed Computing*, 13(1), 2064-2076.
- [66] Garg, P., & Raman, P. K. (2011). Broadcasting Protocol & Routing Characteristics With Wireless ad-hoc networks. *Int. J. Comput. Eng. Manag.*, 12(1), 36-40.
- [67] Garg, P., Arora, N., & Malik, T. (2011). Capacity Improvement of WI-MAX In presence of Different Codes WI-MAX: Speed & Scope of future. *IJCEM*, 12.
- [68] Garg, P., Saroha, K., & Lochab, R. (2011). Review of wireless sensor networks-architecture and applications. *IJCSMS International Journal of Computer Science & Management Studies*, 11(01), 2231-5268.
- [69] Yadav, S., & Garg, P. Development of a New Secure Algorithm for Encryption and Decryption of Images.
- [70] Dixit, A., Sethi, P., & Garg, P. (2022). Rakshak: A Child Identification Software for Recognizing Missing Children Using Machine Learning-Based Speech Clarification. *International Journal of Knowledge-Based Organizations (IJKBO)*, 12(3), 1-15.
- [71] Shukla, N., Garg, P., & Singh, M. (2022). MANET Proactive and Reactive Routing Protocols: A Comparison Study. *International Journal of Knowledge-Based Organizations (IJKBO)*, 12(3), 1-14.
- [72] Arya, A., Garg, P., Vellanki, S., Latha, M., Khan, M. A., & Chhbra, G. (2024). Optimisation Methods Based on Soft Computing for Improving Power System Stability. *Journal of Electrical Systems*, 20(6s), 1051-1058.
- [73] Garg, P. (2025). Cloud security posture management: Tools and techniques. *Technix International Journal for Engineering Research*, 12(3).
- [74] Tyagi, P., Sharma, S., Srivastava, A., Rajput, N. K., Garg, P., & Kumari, M. (2025). AI in Healthcare: Transforming Medicine with Intelligence. In *First Global Conference on AI Research and Emerging Developments (G-CARED 2025)*, New Delhi, India. <https://doi.org/10.63169/GCARED2025.p4>
- [75] Garg, P., Bhatt, M., Parmar, R., & Arsalan, M. (2025). Generative AI: Evolution, Applications, Challenges And Future Prospects. Applications, Challenges And Future Prospects (May 17, 2025).
- [76] Garg, P., Saraswat, P., & Siddiqui, Z. (2025). AI & the Indian Stock Market: A Review of Applications in Investment Decision. <https://doi.org/10.63169/GCARED2025.p10>
- [77] Garg, P., Sharma, S., Mittal, S., Tevatia, R., Tyagi, V. K., & Kapoor, S. (2025). Unlocking Workforce Potential: AI-Powered Predictive Models for Employee Performance Evaluation. <https://doi.org/10.63169/GCARED2025.p21>
- [78] Shrivastava, N., Kalia, A., Roy, R., Sharma, S., Garg, P., & Agarwal, G. (2025). OSINT: A Double-edged Sword. In *First Global Conference on AI Research and Emerging Developments (G-CARED 2025)*, New Delhi, India. <https://doi.org/10.63169/GCARED2025.p22>
- [79] Garg, P., Aditi, A., & Roy, B. (2025). A System of Computer Network: Based On Artificial Intelligence. <https://doi.org/10.63169/GCARED2025.p24>
- [80] Parmar, R., Kapoor, S., Saifi, S., & Garg, P. (2025). Case Study on Intelligent Factory Systems for Improving Productivity and Capability in Industry 4.0 with Generative AI. In *First Global Conference on AI Research and Emerging Developments (G-CARED 2025)*, New Delhi, India. <https://doi.org/10.63169/GCARED2025.p28>
- [81] Singh, R., Sharma, R., Kumar, R., Nafis, A., Siddiqui, M. A. M., & Garg, P. (2025). Detection of Unauthorized Construction using Machine Learning: A Review. In *First Global Conference on AI Research and Emerging Developments (G-CARED 2025)*, New Delhi, India. <https://doi.org/10.63169/GCARED2025.p30>
- [82] Garg, P., Kapoor, S., Singh, V., Sharma, S., & Ankita, A. (2025). A Bridge between Blockchain and Decentralized Applications Web3 and Non-Web3 Crypto Wallets. <https://doi.org/10.63169/GCARED2025.p35>
- [83] Verma, M., Sharma, S., Garg, P., & Singh, A. (2025). The Hidden Dangers of Prototype Pollution: A Comprehensive Detection Framework. In *First Global Conference on AI Research and Emerging Developments (G-CARED 2025)*, New Delhi, India. <https://doi.org/10.63169/GCARED2025.p36>
- [84] Sharma, A., Sharma, S., Garg, P., & Bhardwaj, P. (2025). LockTalk: A Basic Secure Chat Application. In *First Global Conference on AI Research and Emerging Developments (G-CARED 2025)*, New Delhi, India.
- [85] Arora, K., Bawane, R., Gupta, C., Ahmed, K., & Garg, P. (2025). Detection and Prevention of Cyber Attack and Threat using AI. In *First Global Conference on AI Research and Emerging Developments (G-CARED 2025)*, New Delhi, India. <https://doi.org/10.63169/GCARED2025.p38>
- [86] Garg, P., Dhruv, D., Rahman, A. A., Rai, A., Siddiqui, M., & Yadav, D. (2025). Easeviewer: An Esports Production Tool. <https://doi.org/10.63169/GCARED2025.p46>
- [87] Garg, P., Lakshita, L., Mehwish, M., Nazia, N., & Ahmed, K. (2025). Emerging Trend in Computational Technology: Innovations, Applications, and Challenges. Applications, and Challenges (May 17, 2025). <https://doi.org/10.63169/GCARED2025.p51>
- [88] Chauhan, S., Singh, M., & Garg, P. (2021). Rapid Forecasting of Pandemic Outbreak Using Machine Learning. *Enabling Healthcare 4.0 for Pandemics: A Roadmap Using AI, Machine Learning, IoT and Cognitive Technologies*, 59-73.
- [89] Gupta, S., & Garg, P. (2021). An insight review on multimedia forensics technology. *Cyber Crime and Forensic Computing: Modern Principles, Practices, and Algorithms*, 11, 27.
- [90] Shrivastava, P., Agarwal, P., Sharma, K., & Garg, P. (2021). Data leakage detection in Wi-Fi networks. *Cyber Crime and Forensic Computing: Modern Principles, Practices, and Algorithms*, 11, 215.
- [91] Meenakshi, P. G., & Shrivastava, P. (2021). Machine learning for mobile malware analysis. *Cyber Crime and Forensic Computing: Modern Principles, Practices, and Algorithms*, 11, 151.
- [92] Nanwal, J., Garg, P., Sethi, P., & Dixit, A. (2021). Green IoT and Big Data: Succeeding towards Building Smart Cities. In *Green Internet of Things for Smart Cities* (pp. 83-98). CRC Press.
- [93] Gupta, M., Garg, P., & Agarwal, P. (2021). Ant Colony Optimization Technique in Soft Computational Data Research for NP-Hard Problems. In *Artificial Intelligence for a Sustainable Industry 4.0* (pp. 197-211). Springer, Cham.
- [94] Magoo, C., & Garg, P. (2021). Machine Learning Adversarial Attacks: A Survey Beyond. *Machine Learning Techniques and Analytics for Cloud Security*, 271-291.
- [95] Garg, P., Srivastava, A. K., Anas, A., Gupta, B., & Mishra, C. (2023). Pneumonia Detection Through X-Ray Images Using Convolution Neural Network. In *Advancements in Bio-Medical Image Processing and Authentication in Telemedicine* (pp. 201-218). IGI Global.
- [96] Gupta, S., & Garg, P. (2023). 14 Code-based post-quantum cryptographic technique: digital signature. *Quantum-Safe Cryptography Algorithms and Approaches: Impacts of Quantum Computing on Cybersecurity*, 193.
- [97] Prakash, A., Avasthi, S., Kumari, P., & Rawat, M. (2023). PuneetGarg 18 Modern healthcare system: unveiling the possibility of quantum computing in medical and biomedical zones. *Quantum-Safe Cryptography Algorithms and Approaches: Impacts of Quantum Computing on Cybersecurity*, 249.

- [98] Gupta, S., & Garg, P. (2024). Mobile Edge Computing for Decentralized Systems. *Decentralized Systems and Distributed Computing*, 75-88.
- [99] Gupta, M., Garg, P., & Malik, C. (2024). Ensemble learning-based analysis of perinatal disorders in women. In *Artificial Intelligence and Machine Learning for Women's Health Issues* (pp. 91-105). Academic Press.
- [100] Malik, M., Garg, P., & Malik, C. (2024). Artificial intelligence-based prediction of health risks among women during menopause. *Artificial Intelligence and Machine Learning for Women's Health Issues*, 137-150.
- [101] Garg, P. (2024). Prediction of female pregnancy complication using artificial intelligence. In *Artificial Intelligence and Machine Learning for Women's Health Issues* (pp. 17-35). Academic Press.
- [102] Pokhrel, L., Arsalan, M., Rani, P., Garg, P., & Pinheiro, P. R. (2026). AI-Powered Healthcare Solutions: Bridging the Medical Gap in Underserved Communities Worldwide. In *Applied AI and Computational Intelligence in Diagnostics and Decision-Making* (pp. 57-86). IGI Global Scientific Publishing.
- [103] Kapoor, S., Parmar, R., Sharma, N., Garg, P., & Singh, N. J. (2026). AI and Computational Intelligence in Healthcare: An Introductory Guide. In *Applied AI and Computational Intelligence in Diagnostics and Decision-Making* (pp. 1-26). IGI Global Scientific Publishing.
- [104] Pokhrel, L., Kumar, A., Garg, P., Anand, N., & Singh, N. (2026). AI and IoT in Global Health: Ethical Lessons From Pandemic Response. In *Development and Management of Eco-Conscious IoT Medical Devices* (pp. 367-394). IGI Global Scientific Publishing.
- [105] Parmar, R., Singh, A., Garg, P., Sharma, T., & Pinheiro, P. R. (2026). Blockchain for Ethical Supply Chains: Transparency in Medical IoT Manufacturing. In *Development and Management of Eco-Conscious IoT Medical Devices* (pp. 337-366). IGI Global Scientific Publishing.
- [106] Gupta, S., Garg, P., Agarwal, J., Thakur, H. K., & Yadav, S. P. (2024). Federated learning based intelligent systems to handle issues and challenges in IoVs (Part 1). <https://doi.org/10.2174/97898153130311240301>
- [107] Gupta, S., Chaudhary, G., & Garg, P. (2013). Modified AODV Routing Protocol through Cache Memory for Finding New Routing Paths in MANETs. *International Journal of Computer Science & Management Studies*, 13(3).
- [108] Gupta, A., & Garg, P. (2021). Emerging Techniques for Handling Pandemic Challenges. *Enabling Healthcare 4.0 for Pandemics: A Roadmap Using AI, Machine Learning, IoT and Cognitive Technologies*, 189-209.
- [109] Chaudhary, A. P., Mishra, A., Kumar, D., & Garg, P. (2023, April). Human Emotion Recognition using Deep Learning. In *2023 International Conference on Computational Intelligence, Communication Technology and Networking (CICTN)* (pp. 191-197). IEEE.
- [110] Nagpal, S., Garg, P., Gaba, S., & Aggarwal, A. (2023). 13 An improved genetic quantum cryptography model for network communication. *Quantum-Safe Cryptography Algorithms and Approaches: Impacts of Quantum Computing on Cybersecurity*, 177.
- [111] Yadav, M., Swami, V., Kumar, N., & Garg, P. (2025). Comparative study of Repairable Juice Plants using RPGT. *Reliability: Theory & Applications*, 20(2 (84)), 776-783.
- [112] Gupta, A., Garg, P., & Yadav, P. (2025). Role of Generative AI Towards Education and Learning: Present & Future. *TPM-Testing, Psychometrics, Methodology in Applied Psychology*, 32(S6 (2025): Posted 15 Sept), 1059-1076.
- [113] Dalal, P., Beniwal, G., Sharma, V., Garg, P., & Ahmed, K. (2025). Predicting Student Motivation and Engagement through Machine Learning Models. *TPM-Testing, Psychometrics, Methodology in Applied Psychology*, 32(S7 (2025): Posted 10 October), 393-411.
- [114] Gupta, A., Mund, A., Roy, S., Garg, P., & Yadav, D. K. (2025). Trust in AI Systems: A Social-psychological Investigation of Human-AI Collaboration. *TPM-Testing, Psychometrics, Methodology in Applied Psychology*, 32(S7 (2025): Posted 10 October), 428-446.
- [115] Bhardwaj, A., Das, A., Garg, P., & Yadav, S. (2025). Material-Driven Performance Analysis of a Vertical Nanowire Tunnel FET for Analog Applications: Bhardwaj, Das, Garg, and Yadav. *Journal of Electronic Materials*, 1-12.
- [116] Dalal, P., Sharma, B., Sharma, T., Garg, P., & Ahmed, K. (2025). Explainable AI for Understanding Human Decision Making Patterns. *TPM-Testing, Psychometrics, Methodology in Applied Psychology*, 32(S7 (2025): Posted 10 October), 412-427.
- [117] Sharma, K. K., Verma, P. K., Garg, P., & Shrotriya, V. K. (2025, October). Predicting costs and benefits of IoT-based energy management for optimizing sustainable energy storage in rural areas. In *AIP Conference Proceedings (Vol. 3343, No. 1, p. 040017)*. AIP Publishing LLC.
- [118] Ahmed, K., Baranwal, A., Sharma, N., Garg, P., & Singh, N. (2026). The Role of Federated Learning in AI-Powered Integrated Healthcare Solutions. In *Enabling Collaborative Health Intelligence With Federated Learning* (pp. 421-448). IGI Global Scientific Publishing.
- [119] Gupta, S., Garg, P., Agarwal, J., Thakur, H. K., & Yadav, S. P. (2025). Federated learning based intelligent systems to handle issues and challenges in IoVs (Part 2). *Bentham Science Publishers*. <https://doi.org/10.2174/97898153222241250301>
- [120] Garg, P., Pranav, S., & Prerna, A. (2021). Green internet of things (G-IoT): A solution for sustainable technological development. In *Green Internet of Things for Smart Cities* (pp. 23-46). CRC Press.
- [121] Malik, A., Nandal, D., Gupta, V., Garg, P., & Nandal, V. *INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING*.
- [122] Gupta, S., Garg, P., Agarwal, J., Thakur, H. K., & Yadav, S. P. (Eds.). (2025). *Federated learning based intelligent systems to handle issues and challenges in IoVs (Part 2)*.
- [123] Garg, P., Bhatt, M., Parmar, R., & Arsalan, M. (2025). *Generative AI: Evolution, Applications, Challenges And Future Prospects*. Applications, Challenges And Future Prospects (May 17, 2025).
- [124] Kumar, N., Kumar, Y., Khurana, D., Kumar, S., & Garg, P. (2025, November). A Hybrid Ensemble Learning Framework for Interpretable Student Performance Prediction Using Academic and Extracurricular Factors. In *2025 International Conference on Innovations and Emerging Technologies In AI & Communication Systems (IETACS)* (pp. 666-672). IEEE.
- [125] Khurana, D., Kumar, Y., Kumar, N., Kumar, S., & Garg, P. (2025, November). Transformer-Based Movie Recommendation System with Autoencoder-Enhanced Feature Compression. In *2025 International Conference on Innovations and Emerging Technologies In AI & Communication Systems (IETACS)* (pp. 685-690). IEEE.
- [126] Garg, P. (2025, November). Comparative Analysis of Various Neural Networks for Galaxies Classification. In *2025 International Conference on Innovations and Emerging Technologies In AI & Communication Systems (IETACS)* (pp. 697-701). IEEE.
- [127] Saggi, A. K., Babbar, N., & Garg, P. (2025, November). Health-Guard AI: Integrated Health Report Management and Analysis. In *2025 International Conference on Innovations and Emerging Technologies In AI & Communication Systems (IETACS)* (pp. 614-623). IEEE.
- [128] Kumar, S., Kumar, Y., Kumar, N., Khurana, D., & Garg, P. (2025, November). Hybrid FCM-DNN Model for Uncertainty-Aware Air Quality Classification Using Multi-Pollutant Data. In *2025 International Conference on Innovations and Emerging Technologies In AI & Communication Systems (IETACS)* (pp. 679-684). IEEE.
- [129] Babbar, N., Singh, H. V., Bendale, S., & Garg, P. (2025, November). Stock Market Price Prediction Using Big Data Analysis: A Performance Evaluation Study. In *2025 3rd International Conference on Computational Intelligence and Network Systems (CINS)* (pp. 1-6). IEEE.

- [130] Singh, A. K., Kori, G., Garg, P., & Srivastava, G. (2025, November). Bank Churn Prediction Using Machine Learning. In 2025 IEEE 7th International Conference on Computing, Communication and Automation (ICCCA) (pp. 1-6). IEEE.
- [131] Bhardwaj, A., Das, A., Garg, P., & Yadav, S. (2026). Material-Driven Performance Analysis of a Vertical Nanowire Tunnel FET for Analog Applications. *Journal of Electronic Materials*, 55(1), 1099-1110.
- [132] Srivastava, A. K., Shankdhar, D., Ror, R., & Garg, P. (2026). Harnessing YOLOv5 for real-time object detection: A cloud-based approach. In *Recent Advances in Computational Methods in Science and Technology* (pp. 441-450). CRC Press.
- [133] Srivastava, A. K., Shukla, A., Gupta, H., Saxena, K., & Garg, P. (2026). Towards intelligent attendance management system with face recognition using LBPH algorithm. In *Recent Advances in Computational Methods in Science and Technology* (pp. 8-15). CRC Press.
- [134] Srivastava, A. K., Garg, P., & Pandey, H. (2026). Vedcure: Towards intelligent ayurvedic drug recommendation and disease prediction. In *Recent Advances in Computational Methods in Science and Technology* (pp. 16-23). CRC Press.
- [135] Upadhyay, D., Garg, P., & Babbar, N. (2026). Blockchain and IoT based smart contract framework for efficient and secure product life management. *Discover Internet of Things*.
- [136] Singh, A., Parmar, R., Bhardwaj, P., Sharma, V., & Garg, P. (2026). Fusion of Aerial Networks with Advanced Computing Paradigms. *Edge Computing and Aerial Platforms*, 355-367.
- [137] Kumari, M., Baranwal, A., Sonal, & Garg, P. (2026). Application of Aerial Edge Computing in Disaster Management. *Edge Computing and Aerial Platforms*, 103-122.
- [138] Aditi, Saraswat, P., Sharma, V., & Garg, P. (2026). Advances in Aerial Platforms and Edge Computing. *Edge Computing and Aerial Platforms*, 123-143.
- [139] Garg, P., Arora, K., Bawane, R., Gupta, C., & Ahmed, K. (2025). Detection and Prevention of Cyber Attack and Threat using AI.
- [140] Ahmed, K., Ahmed, K., Khan, J., Garg, P., Seth, S., & Mallik, S. (2025). Principal Component Analysis Based Clustering of Insecticides and Molecular Docking of Pyrethroids Insecticides.
- [141] Kumar, B., Kumar, A., Nanwal, J., Garg, P., & Patnaik, P. (2025, November). Ensemble of YOLOv5 and Segment Anything Model for Brain Tumor Detection. In 2025 2nd International Conference on Advanced Computing and Emerging Technologies (ACET) (pp. 1-5). IEEE.
- [142] Arsalan, M., Anas, M., & Garg, P. (2025). Transparent AI for Drug Discovery and Development. Available at SSRN 5844242.
- [143] Singh, A., Bhardwaj, P., Garg, P., & Singh, N. (2026). Introduction to explainable artificial intelligence in healthcare. In *Explainable AI in Clinical Practice* (pp. 23-44). Academic Press.
- [144] Kapoor, S., Singh, A., Garg, P., & Ramasamy, L. K. (2026). Explainable artificial intelligence in a diagnostic support system. In *Explainable AI in Clinical Practice* (pp. 131-145). Academic Press.
- [145] Ahmed, K., Anas, M., & Garg, P. (2026). Case studies on unlocking the potential of Industry 4.0 for sustainable manufacturing through generative AI-driven innovations. Available at SSRN 6356958.
- [146] Garg, P., & Oruganti, S. K. (2026, March). AI Assisted Routing Optimization in Opportunistic IoT Networks using Machine Learning: A Comprehensive review on Protocols & Simulators. In *Sustainable Global Societies Initiative* (Vol. 1, No. 4). Vibrasphere Technologies.
- [147] Arsalan, M., Pokhrel, L., & Garg, P. (2026). Architecture, Components, and tools in Integrated AI-Augmented Intelligence: A design perspective. Components, and tools in Integrated AI-Augmented Intelligence: A design perspective (March 19, 2026).
- [148] Singh, H., Ahmed, K., & Garg, P. (2026). Human Versus Machine Customer Behaviour and Functional Differences. Available at SSRN 6441098.
- [149] Saraswat, P., & Garg, P. (2026). Soft Computing In AI Agents.
- [150] Saraswat, P., & Garg, P. (2026). Water Quality Prediction Using IOT Sensors and Deep Networks.
- [151] Arsalan, M., Ahmed, K., & Garg, P. (2026). Machine learning for Anomaly detection in sensor networks. Available at SSRN 6441518.
- [152] Kumari, M., & Garg, P. (2026). Hybrid Cloud Infrastructure: Models, Benefits and Security And Challenges. Benefits and Security and Challenges (March 19, 2026).
- [153] Singh, H., & Garg, P. (2026). Demystifying Artificial Distributed Intelligence (ADI). Available at SSRN 6442698.
- [154] Saraswat, P., & Garg, P. (2026). Human AI Collaboration the future of clinical decision making.
- [155] Saraswat, P., & Garg, P. (2026). Breaking Data Boundaries Federated Learning in Digital Healthcare.
- [156] Singh, A., Parmar, R., Bhardwaj, P., Sharma, V., & Garg, P. (2026). Fusion of Aerial Networks with Advanced Computing Paradigms. *Edge Computing and Aerial Platforms*, 355-367.
- [157] Kumari, M., Baranwal, A., Sonal, & Garg, P. (2026). Application of Aerial Edge Computing in Disaster Management. *Edge Computing and Aerial Platforms*, 103-122.
- [158] Aditi, Saraswat, P., Sharma, V., & Garg, P. (2026). Advances in Aerial Platforms and Edge Computing. *Edge Computing and Aerial Platforms*, 123-143.
- [159] Garg, P. (2026). Zero-Trust Security Enforcement through AI-Powered Anomaly Detection in Cloud Systems. *Journal of Artificial Intelligence in Governance and Public Policy (JAIGPP)*, 1(1), 1-8.
- [160] Singh, A. P., Sharma, A., & Garg, P. (2026, January). AI-Powered Adaptive Mock Interview Generation System. In 2026 International Conference on AI-Driven Smart Systems and Ubiquitous Computing (ICAUC) (pp. 1421-1426). IEEE.
- [161] Raghav, A., Mishra, A., & Garg, P. (2026, January). Enhancing Healthcare Access: An AI-driven Chatbot for Doctor Appointment Management. In 2026 International Conference on AI-Driven Smart Systems and Ubiquitous Computing (ICAUC) (pp. 910-914). IEEE.
- [162] Kumar, B., Chauhan, D., Singh, H., Verma, H., Sahu, K., & Garg, P. (2026, January). Decoding Linear A with Artificial Intelligence: A Comprehensive Machine Learning and NLP Framework. In 2026 International Conference on AI-Driven Smart Systems and Ubiquitous Computing (ICAUC) (pp. 1-7). IEEE.
- [163] Gupta, V., Chakravarti, L., Akhtar, M. M., Maheshwari, P., Garg, P., & Tiwari, D. (2026, January). A Sentence-Level Risk Estimator for Identifying Hallucinations in Generative AI. In 2026 International Conference on AI-Driven Smart Systems and Ubiquitous Computing (ICAUC) (pp. 1619-1626). IEEE.
- [164] Patnaik, P., & Garg, P. (2026). Principles of Artificial Intelligence: Cognitive Architectures, Large Language Models, and Computational Limits. Deep Science Publishing.
- [165] Singh, K., & Garg, P. (2026). Trustworthy Deep Learning: Robustness, Uncertainty Quantification, and Adversarial Resilience. Deep Science Publishing.
- [166] Garg, P. (2026). Survey of Load Balancing Strategies in Fog-Cloud Architectures for IoT Integration. *International Journal of Research Publications in Engineering, Technology and Management (IJRPETM)*, 9(2), 595-604.
- [167] Tiwari, D., Bhati, B. S., Garg, P., & Babbar, N. (2026). A novel lexicon dictionary and CNN-LSTM employed hybrid approach for sentiment detection of COVID-19 vaccines. *Scientific Reports*.

- [168] Sethi, P., Kumar, B., Garg, P., & Yadav, P. (2026, February). Graph Neural Networks for Causal Inference in Climate Science: A Novel Approach to Model Modeling Complex Interactions. In 2026 International Conference on Intelligent Computing and Automation for Sustainable Solutions (ICASS) (pp. 1-7). IEEE.
- [169] Kumar, B., Garg, P., Garg, A., Nanwal, J., & Yadav, P. (2026, February). An AI-Centric Multimodal Framework for Cognitive Productivity and Digital Wellbeing. In 2026 International Conference on Intelligent Computing and Automation for Sustainable Solutions (ICASS) (pp. 1-7). IEEE.
- [170] Ahmed, K., Mathur, T., Chaudhary, Y., Anshika, A., Garg, P., & Tiwari, D. (2026, April). ParseRAG: Layout-Aware Open-Source Alternative. In 2026 Eight International Conference on Computational Intelligence and Communication Technologies (CCICT) (pp. 799-804). BM Institute of Engineering & Technology.
- [171] Yadav, P., Raj, M., Kaushik, S., Goel, P., Garg, P., & Tiwari, D. (2026, April). Employing Enhanced LSTM and Large Language Models for Stock Price Prediction Using Sentiment Analysis. In 2026 Eight International Conference on Computational Intelligence and Communication Technologies (CCICT) (pp. 617-622). BM Institute of Engineering & Technology.
- [172] Gupta, V., Sharma, H., Nagpal, N., Garg, P., & Tiwari, D. (2026). Feel Frames: An Emotion Based Video Summarization using Multimodal Deep Learning. *Procedia Computer Science*, 282, 715-724.
- [173] Singh, H., Ahmed, K., Kumar, U., & Garg, P. (2026, July). Comparative Analysis of Machine Learning and Large Language Model Architectures for Intelligent Medical Diagnosis and Clinical Decision Support Systems. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 125-138).
- [174] Singh, H., Ahmed, K., Prajapati, P., & Garg, P. (2026, July). Federated Learning with Differential Privacy: A Comprehensive Framework for Privacy-Preserving Distributed Machine Learning. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 139-155).
- [175] Singh, H., Ahmed, K., & Garg, P. (2026, July). Survey on Internet of Things (IOT) Forensics: Challenges, Approaches, Privacy-Aware Frameworks and Open Issues. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 156-170).
- [176] Singh, H., Ahmed, K., & Garg, P. (2026, July). Supply Network Disruption Prediction using Heterogeneous Graph Neural Networks and Temporal Embeddings. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 171-186).
- [177] Saraswat, P., Kumari, S., & Garg, P. (2026, July). Future Directions in Digital Social Linguistics. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 187-203).
- [178] Saraswat, P., Yadav, A., & Garg, P. (2026, July). Comparing Supervised and Unsupervised Learning Approaches for Specific Tasks. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 220-236).
- [179] Saraswat, P., Aishwarya, S., & Garg, P. (2026, July). Foundation of Self-Learning AI for Drones. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 237-254).
- [180] Ahmed, K., Singh, H., Anas, M., & Garg, P. (2026, July). The Future of Artificial Intelligence in Healthcare. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 255-270).
- [181] Ahmed, K., & Garg, P. (2026, July). AI-DRIVEN FAULT DETECTION AND DIAGNOSTICS. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 271-292).
- [182] Kumari, M., Pandey, S., & Garg, P. (2026, July). Quantum AI for Drug Discovery and Material Science. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1).
- [183] Kumari, M., Jha, S. K., & Garg, P. (2026, July). Opportunities and Challenges of AI in Education. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 328-341).
- [184] Pokhrel, L., Kumar, A., & Garg, P. (2026, July). Impact of AI on Employment and Workforce Transformation. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 342-356).
- [185] Singh, H., Ahmed, K., Kumar, A., & Garg, P. (2026, July). AI-Driven Optimization in Supply Chain Operations: A Machine Learning Framework for Demand Forecasting and Inventory Management. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 357-371).
- [186] Saraswat, P., Kumar, R., & Garg, P. (2026, July). Human Perception and Virtual Reality: What Is the Connection?. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 409-427).
- [187] Ahmed, K., Arsalan, M., Sharma, P. K., & Garg, P. (2026, July). Quantum Technology in Medical Diagnosis. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 428-448).
- [188] Saraswat, P., Sharma, P., & Garg, P. (2026, July). The Challenges of AI Implementation in Business. In *Adroid Conference Series: Engineering and Technology* (Vol. 2, No. 1, pp. 204-219).
- [189] Trivedi, S., Beri, T., Arora, A., & Garg, P. (2026, June). NoduleNet: Deep Learning-Based Computational Model for Thyroid Cancer Detection. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-7). IEEE.
- [190] Joshi, V., Tyagi, U., Jakhar, V., Gaur, M., & Garg, P. (2026, June). HealthQ: A Retrieval-Augmented Health Policy Query Chatbot. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.
- [191] Saraswat, P., Sharma, V., & Garg, P. (2026, June). Energy Optimization In Smart IOT Enabled Smart Home System Through Machine Learning Techniques. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.
- [192] Saraswat, P., Tiwari, B., & Garg, P. (2026, June). An Empirical Study on Salary Determinants Using Machine Learning. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-5). IEEE.
- [193] Pokhrel, L., Arsalan, M., Tiwari, B., & Garg, P. (2026, June). Food Delivery Time Prediction Using Machine Learning Algorithms: A Comparative Study with Hybrid XGBoost-LSTM Architecture. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-7). IEEE.
- [194] Garg, P., & Kumar, V. (2026, June). Millimeter-Wave 60 GHz On-Chip Dipole Antenna with Silicon Lens Integration in SiGe Technology. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-5). IEEE.
- [195] Arora, A., Trivedi, S., Tyagi, R., & Garg, P. (2026, June). LungVision: Classifying lung diseases using CNN ensembles. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.
- [196] Arora, A., Trivedi, S., Sharma, R., Tiwari, R., Chaudhary, S., & Garg, P. (2026, June). Comparative Study of ResNet-152 and ViTPose-G for Human Pose Detection on the COCO Yoga Dataset. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.
- [197] Kaur, M., Chhillar, V., Kumar, V., & Garg, P. (2026, June). Exploration and Recognition of Autism Spectrum Disorder through Machine Learning Approaches. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.

- [198] Singh, R., Singh, G., Kumar, V., Kumari, M., Kumar, A., & Garg, P. (2026, June). Optimization of Battery Electric Vehicle Thermal Management System Using MATLAB Simscape for Enhanced Charging Efficiency. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.
- [199] Chauhan, A., Singh, G., Kumar, V., Kumari, M., Kumar, A., & Garg, P. (2026, June). An Intelligent Thermal Control and Optimization Framework for Efficient and Safe Battery Charging in Electric Vehicles. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.
- [200] Kumar, S., Kumar, Y., Khurana, D., Kumar, N., & Garg, P. (2026, June). AI-Driven Discovery of Cybercrime Insights from Diverse Web Data. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.
- [201] Chaudhary, N., Tyagi, N., Sharma, N., Gupta, M., & Garg, P. (2026, June). BusOptima-Bus Scheduling & Route Management System. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.
- [202] Srivastava, A. K., Singh, A., Umrao, A., & Garg, P. (2026, June). A Retrieval-Augmented Framework for Context-Aware Document Querying. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-7). IEEE.
- [203] Gupta, T., Bhargava, D., & Garg, P. (2026, June). A Comprehensive Survey on Malware Botnet Detection Using Deep Learning in IoT and Android Platforms with Edge AI. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-7). IEEE.
- [204] Trivedi, S., Shukla, B., Garg, Y., Satish, T., & Garg, P. (2026, June). Enhancing Social Media Security Through Explainable Ensemble Learning. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.
- [205] Arora, P., Arora, R., Tiwari, D., & Garg, P. (2026, June). Automated Brain Tumor Classification from MRI Images Through InceptionNet Architecture. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-6). IEEE.
- [206] Srivastava, S., Singh, S., Goswami, S., Gupta, M., & Garg, P. (2026, June). SignSpeak: A Real-Time Deep Learning Framework for Bidirectional Sign Language Translation. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-7). IEEE.
- [207] Srivastava, A. K., Garg, P., & Singh, A. (2026, June). ExplainFusion: An Ensemble Explainable AI Framework with Evidence-Grounded LLM-Based Clinical Explanations. In 2026 6th International Conference on Emerging VLSI and Semiconductor Technology for AI and Computing Applications (EVST) (pp. 1-8). IEEE.