iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** X **Month of publication:** October 2026

DOI: https://doi.org/10.22214/ijraset.2026.85474

www.ijraset.com

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

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

A Comparative Analysis of Pose Estimation Models and a Proposed Media Pipe-Based Framework for Vision-Based Fall Detection

Rohan Singh, Nupur Thakkar, Vedant Lade, Siddhesh Bhiwapurkar
Department of Computer Engineering, St.Vincent Pallotti College of Engineering and Technology, Nagpur, India

Abstract: ***Vision-based fall detection increasingly relies on human pose estimation to convert raw video into structured skeletal data for posture and motion analysis. This paper presents a literature-based comparative study of five widely used pose estimation models — MediaPipe Pose Landmarker, OpenPose, AlphaPose, MoveNet, and HRNet — examined across architecture, accuracy, speed, computational cost, edge-deployment readiness, multi-person capability, and healthcare suitability. The comparison surfaces a research gap: existing fall-detection literature rarely justifies its choice of pose-estimation backbone through systematic comparison, and almost no published study evaluates these backbones together under fall-specific, resource-constrained conditions. Motivated by this gap, and by published evidence favoring lightweight, CPU-feasible architectures for continuous healthcare monitoring, this paper proposes a conceptual prototype framework for fall detection built around MediaPipe Pose Landmarker, combining pose-based posture analysis with temporal inactivity reasoning to suppressfalsealarms.Theproposedframeworkispresentedstrictlyatthearchitecturalandworkflowlevel;noexperimentalimplementation, datasetevaluation,orperformanceresultsarereportedhere.Thispaperestablishestheliterature-grounded foundation and prototype design that will be implemented and experimentally validated in a subsequent study.***

Keywords: ***Humanposeestimation;falldetection;MediaPipe;BlazePose;edgecomputing;healthcaremonitoring;comparativeanalysis; prototype framework***

I. INTRODUCTION

Fallsremainamongtheleadingcausesofunintentionalinjuryin older adults, and detection latency, not just detection accuracy, often determines outcome severity [4]. Fall detection technologies fall into three broad families: wearable sensors, ambient sensors, and vision-based systems. Vision-based approaches are attractive because they require no user compliance and, when reasoned over a skeletal representation rather than raw pixels, can offer a degree of privacy by transmitting only joint coordinates rather than video [1].

Skeleton-based fall detection depends on an upstream human pose estimation model to convert each frame into body keypoints. Several architectures are available for this role — OpenPose, AlphaPose, MoveNet, HRNet, and MediaPipe'sBlazePose-based Pose Landmarker — each reflecting different priorities around keypoint density, multi-person handling, and on-deviceefficiency.Areviewofrecentfall-detectionliterature shows that the choiceamongthese backbones is rarelyjustified through systematic comparison; teams typically adopt whichevertoolisfamiliarorconvenient,withlittlediscussionof alternatives.

This paper addressesthat gapintwostages. First,it synthesizes published literature comparing the five pose-estimation architectureslistedaboveacrosstheparametersmostrelevantto continuous, healthcare-oriented deployment. Second, building onthissynthesisandonourprojectsynopsisforanAI-basedfall detectionsystem,itproposesaconceptualprototypeframework centeredon MediaPipe PoseLandmarker.Thisisthefirstoftwo plannedpapers:thepresent paperisstrictlyliterature-basedand architectural, while a second paper, to follow, will report the implementation and experimental validation of the proposed framework.

The contribution of this paper is fourfold: (i) a concise comparativeanalysisoffivepose-estimationmodelsforfalldetection; (ii) identification of a research gap concerning backbone justification and fall-specific evaluation; (iii) a literature-supported justification for selecting MediaPipe Pose Landmarker;and(iv)aproposedprototypearchitecturederived from this literature and from our project synopsis, intended to guide future implementation.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue X Oct 2026- Available at www.ijraset.com

II. RELATED WORK

General surveys of human pose estimation [10], [11] catalogue thefield'sarchitecturalevolutionandconsistentlyflagocclusion and crowded scenes as unresolved challenges, but none addressesfall detectionspecifically.Theonlycontrolled,same-hardware comparison available, Jo and Kim [6], measured OpenPose, PoseNet, and two MoveNet variants on identical mobile hardware, finding MoveNet Lightning fastest and OpenPose the only model capable of multi-person estimation; this study did not include MediaPipe's Pose Landmarker or AlphaPose.

Withinthefall-detectionliterature,ahandfulofbackbonesrecur. OpenPose was used in some of the earliest pose-based fall pipelines [12]. MediaPipe Pose has become common in CPU-constrained deployments [13]–[16], with Sykes [15] explicitly comparingOpenPose, PoseNet, and MediaPipebeforefavoring a lightweight backbone for Raspberry-Pi-class hardware. AlphaPoseappearsmostoftenpairedwithspatio-temporalgraph convolutional classifiers, reporting strong accuracy on UR Fall and similar datasets [17], [18]. Notably, justification for backbone choice in this literature is almost always a single sentence citing computational convenience rather than a structured comparison — a pattern this paper aims to address directly.

III. COMPARATIVE ANALYSIS OF POSE ESTIMATION MODELS

Fig.1situatesthefivecomparedarchitecturesalongtheirrelease timeline,fromearlyheatmap-basednetworksthroughtoday's mobile-optimized models. Each is summarized below and compared in Table I.

MediaPipePoseLandmarker(BlazePose)[1]usesalightweight detector-tracker CNN producing 33 keypoints and reports over 30 FPS on a mobile CPU. It is single-person bydefault but has accumulatedthe broadest healthcare-adjacent validationamong thefivemodels,includinggaitanalysisagainstamotion-capture reference [22] and post-stroke kinematic tracking [23].

OpenPose [2] was the first real-time multi-person 2D pose estimator, using bottom-up Part Affinity Fields. It remains an accuracy reference point but is roughly an order of magnitude slowerthanlightweightalternativesonequivalenthardware[6], limiting CPU-only deployment.

AlphaPose [3] is atop-down system that corrects for imprecise detector bounding boxes via a Symmetric Spatial Transformer Network, and its whole-body variant adds identity-aware tracking. It pairs well with graph-based temporal classifiers for fall recognition [17], [18], but its accuracy is bound by the upstream person detector, which can fail outright on fall-like, motion-blurred frames.

MoveNet[7]isasingle-stage,CenterNet-stylebottom-upmodel builtforTensorFlowLite.Itisthefastestmodelindirectmobile comparison [6], but its 17-point topology and single-person design have limited its adoption in fall-detection-specific research relative to MediaPipe and OpenPose.

HRNet [5] maintains high-resolution feature representations throughout the network, yielding state-of-the-art keypoint accuracy on COCO at a high memory and compute cost. It functions chiefly as an accuracy reference point rather than a deployment candidate for constrained hardware [10], [11].

As Table I summarizes, the five models trade off along broadly the same axes: OpenPose and AlphaPose offer native or near-native multi-person capability and richer keypoint vocabularies at a higher computational cost; HRNet offers the strongest raw keypointaccuracyatthehighestcost;MoveNetoffersthefastest inference but the coarsest topology and weakest multi-person support; and MediaPipe Pose Landmarker occupies a middle position on raw speed and accuracy while leading the group on accumulated healthcare-validation evidence and CPU-only deployability. Table IIsummarizesrepresentativefall-detection studies built on these backbones, illustrating how each has actually been used in practice rather than how it performs on generic benchmarks alone.

IV. RESEARCH GAP

Fig.4illustratesthegapconceptually:pose-estimationaccuracy research and fall-detection deployment literature overlap only partially, and the overlap rarely includes a systematic, same-hardware backbone comparison.

Three shortcomings follow directly from Sections II–III. First, fall-detection papers seldomjustify their pose-backbone choice comparatively,relyinginsteadonconventionorconvenience[12]–[16]. Second, the accuracy figures most commonly cited for these models reflect COCO/MPII-style benchmarks dominatedbyupright,well-litposes,nottheextremeorientationchanges and motion blur characteristic of an actual fall. Third, computational-cost reporting is inconsistent across studies, and no published comparison evaluates MediaPipe Pose Landmarker, OpenPose, AlphaPose, MoveNet, and HRNet together on identical hardware for a fall-relevant task.

©IJRASET: All Rights are Reserved

V. PROPOSED PROTO TYPE FRAMEWORK

Building on the comparative analysis of Section III and the research gap of Section IV, and consistent with our project synopsis for an AI-based fall detection system, this section presents a conceptual prototype framework. The framework is describedatthearchitecturalandworkflowlevelonly;ithasnot yet been implemented or experimentally evaluated, and no performance figures are claimed for it.

A. System Overview

Fig. 2 shows the overall proposed architecture. A camera supplies continuous video to a pipeline that runs entirely on a local edge device — a laptop or embedded board with no mandatorydiscreteGPU—consistentwiththehardwarebudget definedin our synopsis. Frames are acquiredand preprocessed, passed to MediaPipe Pose Landmarker for keypoint extraction, and the resulting skeleton drives feature extraction, hybridfall-detection logic, and, when a fall is confirmed, alert generation toward a caregiver or notification endpoint. Running the pose-estimation and decision stages locally avoids continuous transmission of raw video, addressing the privacy concerns noted in Section I.

B. WorkflowandModuleInteraction

Fig. 3 details the intended workflow. Each captured frame is preprocessed and passedthrough MediaPipe Pose Landmarker; if nokeypoints are detected, thesystem simply continues tothe next frame.Whenkeypointsareavailable,featuressuchasjoint angles,trunkorientation,andcentroidpositionareextractedand checked for a posture change consistent with a fall. Rather than triggering an alert from posture alone, the framework evaluates the duration of subsequent inactivity, distinguishing an uncontrolledfall from adeliberatesit orlie-down —thehybrid logic motivated by the false-positive concerns identified in Section IV. Only when inactivity exceeds a configurable threshold is the fall confirmed and an alert generated.

Fig. 6 decomposes this workflow into interacting software modules: a Video Capture Module feeding a Preprocessing Module, a Pose Estimation Module (MediaPipe), a Feature Extraction Module, a Posture Analysis Module, a Temporal/Inactivity Module, a False-Positive Reduction Module, and an Alert Generation Module. The False-Positive Reduction Module feeds back into Posture Analysis, allowing borderlinecases—bending,crouching,sittingabruptly—tobe re-evaluated against accumulated temporal context rather than classified from a single frame.

C. Literature-SupportedModelSelection

Fig. 5 summarizes, as a decision flow, how the comparative analysis of Section III informs backbone selection for this framework. Where densemulti-persontrackingisthedominant requirement, AlphaPose or OpenPose remain the more defensiblechoicegiventheirnativemulti-personhandling;whereadedicatedGPUisguaranteed,HRNet-classmodelsoffer the strongest raw accuracy. For the deployment envelope targetedbyoursynopsis—continuous,CPU-onlymonitoringof a typically single-occupant or sparsely-occupied room — the literaturereviewedinSectionIIIinsteadfavorsMediaPipePose Landmarker, on the strength of its reported real-time CPU performance [1], its accumulating clinical-adjacent validation [22], [23], and its consistent adoption across recent CPU-constrained fall-detection systems [13]–[16].

D. RelationtotheIdentifiedResearchGap

TheproposedframeworkaddressesthegapsofSectionIVatthe design stage, ahead of implementation. The backbone-justificationgapisaddressedbySectionIIIandFig.5ratherthan by convention. The fall-specific-accuracy gap is addressed by combining posture analysis with temporal inactivity reasoning (Fig. 3) rather thanrelyingonraw per-frame keypoint accuracy alone. The same-hardware-comparison gap is not resolved by this paper — it remains explicitly future work, to be addressed once the framework described here is implemented and benchmarked in our second paper.

VI. CONCLUSION

This paper compared five pose-estimation architectures for vision-based fall detection — MediaPipe Pose Landmarker, OpenPose, AlphaPose, MoveNet, and HRNet — and identified aresearchgapconcerningsystematicbackbonejustificationand fall-specific, same-hardware evaluation. Building on this analysis and on our project synopsis, a conceptual prototype framework centered on MediaPipe Pose Landmarker was proposed, combining posture analysis with temporal inactivity reasoning within a CPU-only, edge-deployable architecture.

This conclusion is intentionally bounded: the proposed framework has not been implemented or experimentally validated,andnoaccuracy,latency,orotherperformancefigures areclaimedinthispaper.Futurework,tobereportedinasecond paper, involves implementing the framework described in Section V, evaluating it experimentally on fall-detection benchmarks,andreportingprecision,recall,andlatencyresults— together with a same-hardware comparison against the alternative backbones discussed in Section III.



Fig.1.Evolutionofhumanposeestimationmodelsrelevanttothisstudy, fromearlyheatmap-basednetworkstomobile-optimized architectures.

TABLEI.COMPARISON OFPOSEESTIMATION MODELS

Model	Accuracy	Speed	Comp.Cost	Edge Deployment	Multi-Person	HealthcareSuitability
MediaPipe(BlazePose)	Good	Veryhigh	Low	Strong(purpose-built)	Single (default)	High—broadestclinical validation
OpenPose	High(benchmarkref.)	Low(~12xslower[6])	High	Weak(GPU-bound)	Native,multi-person	Moderate
AlphaPose	High (in ST-GCNpipelines)	Moderate(GPU)	High	Weak	Native,whole-body	Emerging(fall-specific)
MoveNet	Lower(raw,[6])	Highest[6]	Low	Strong(TFLite-native)	Single(released ver.)	Limited literature
HRNet	Highest(COCO)	Low	Veryhigh	Weak	Dependsonhead	Limited(research-stage)

TABLE II.SUMMARYOFREPRESENTATIVELITERATURE

Study	Backbone	Application	Key Finding
Jo&Kim[6]	OpenPose/PoseNet/MoveNet	Mobilecomparison	MoveNetfastest;OpenPoseonlymulti-person
Wang&Deng[13]	MediaPipe	Elderlyfalldetection	~90%accuracyonCPU-only laptop
Sykes[15]	OpenPose/PoseNet/MediaPipe	Low-powerfalldetection	98%accuracyonRaspberry-Pi-classdevice
Shinet al.[17]	AlphaPose+ST-GCN	Fall recognition	Upto98.97%onUR Fall;someframesdiscarded
Hiietal.[22]	MediaPipe	Gaitanalysis	Good-to-excellentagreementwith Vicon

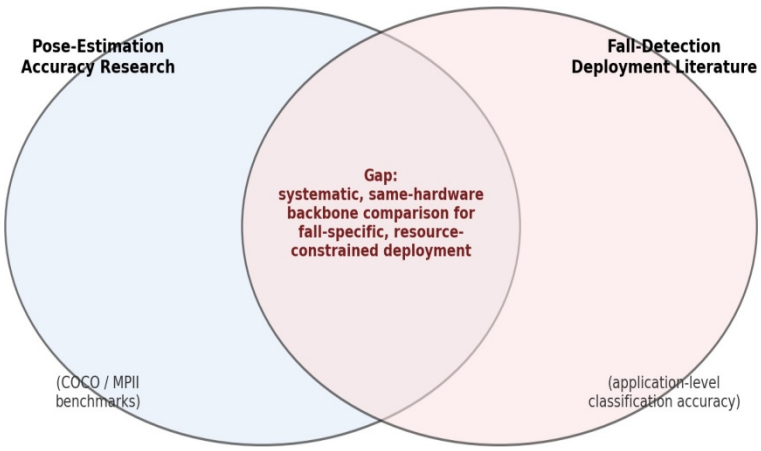


Fig.4.Conceptualillustrationoftheresearchgapbetweenpose-estimationaccuracyresearchandfall-detectiondeployment literature

TABLEIII.RESEARCHGAPANALYSIS

IdentifiedGap	Description	AddressedbyProposedFramework
Backbonejustification	Pose-backbonechoicerarelycomparedsystematically	Justifiedvialiteraturesynthesis(Sec.III,VI)
Fall-specificaccuracy	COCO/MPIIaccuracydoesnotreflectfallpostures	Hybridposture+temporallogic,notrawkeypointaccuracyalone
Inconsistentcostreporting	FPS/latencyreportednon-uniformlyacrossstudies	FrameworkdesignedforCPU-onlyoperationby construction
Same-hardwarecomparison	Nostudycomparesallfivebackbones together	Identifiedexplicitlyasfutureimplementationwork

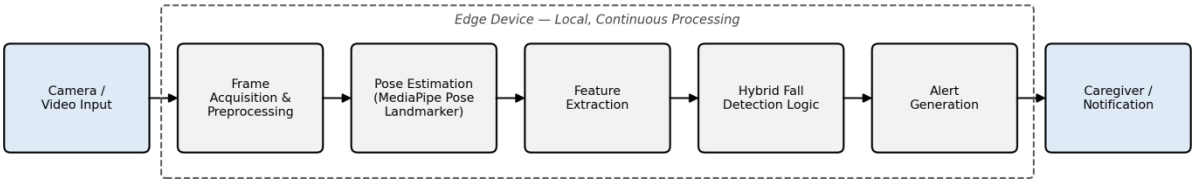


Fig.2.Overallproposedprototypearchitecture,showingtheedge-resident processingpipelinefromvideocapturetoalertgeneration

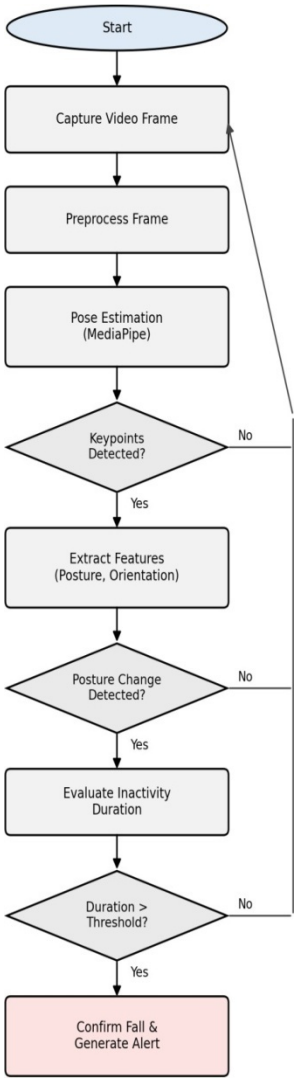


Fig.3.Workflowoftheproposedfalldetectionframework,combiningpostureanalysiswithtemporalinactivity reasoning.

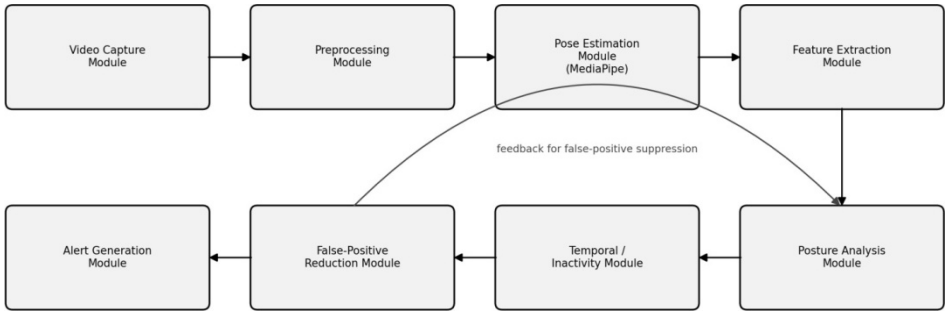


Fig.6.Moduleinteractiondiagramfortheproposedframework, showingdataflowandthefalse-positivefeedback path

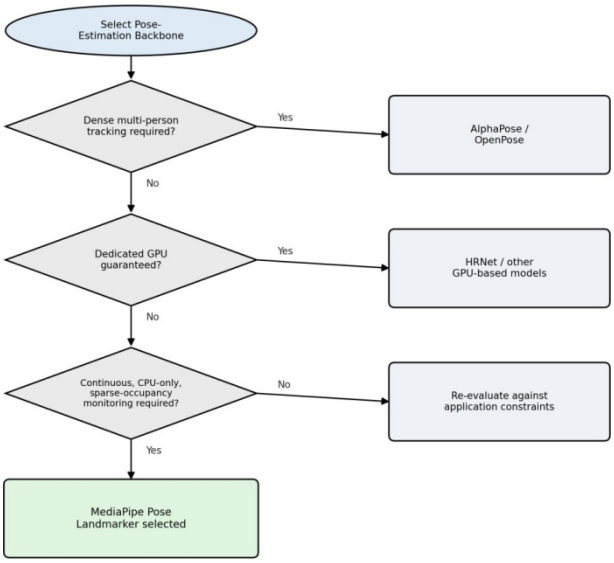


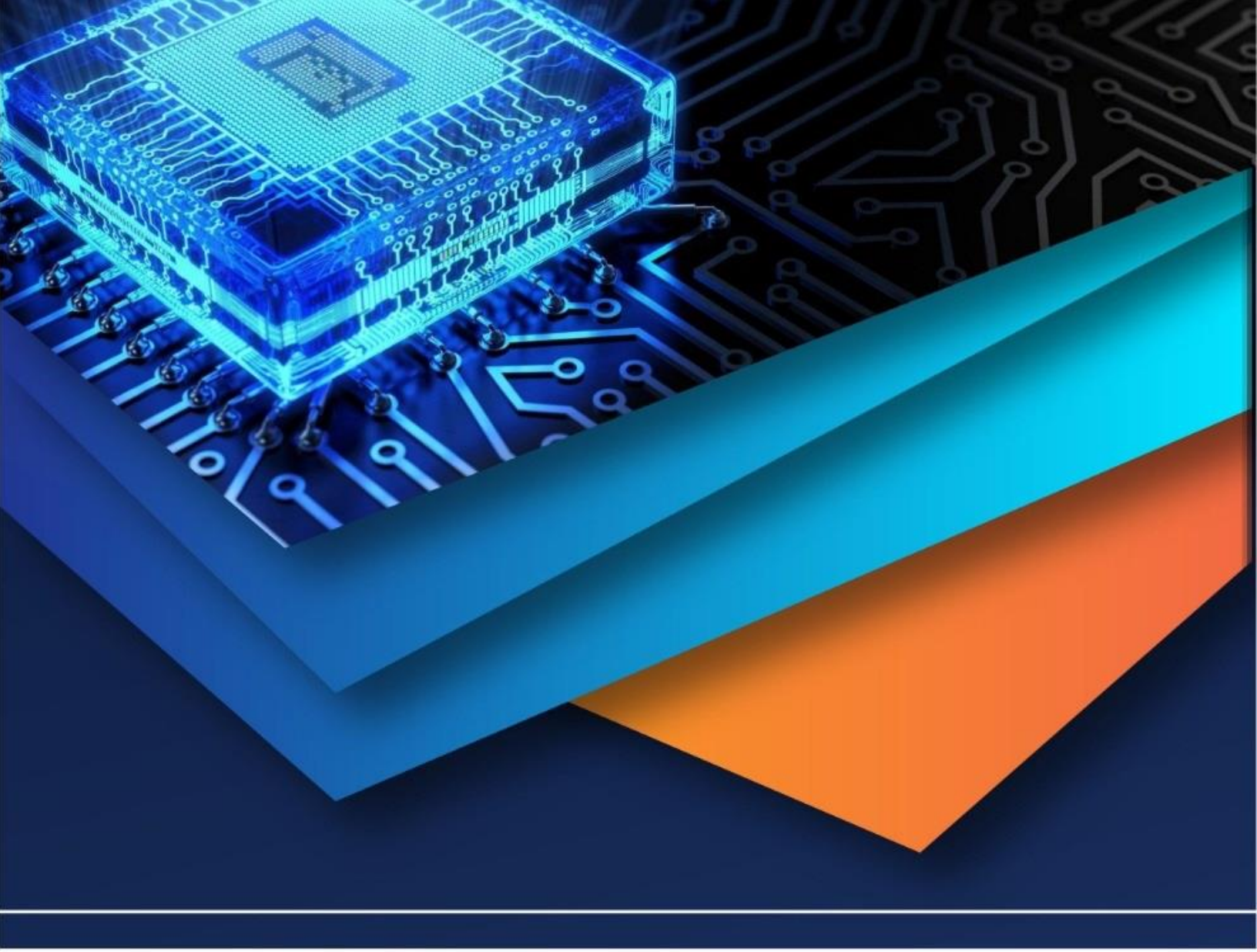
Fig.5.Decisionflow showinghowapplicationconstraintsleadtotheliterature-supportedselectionofMediaPipePoseLandmarker

REFERENCES

[1] V.Bazarevsky,I.Grishchenko,K.Raveendran,T.Zhu,F.Zhang,andM.Grundmann,"BlazePose:On-devicereal timebodypose tracking,"arXiv:2006.10204,2020.
[2] Z.Cao,G.Hidalgo,T.Simon,S.E.Wei,andY.Sheikh,"OpenPose:Realtimemultiperson2Dposeestimationusingpartaffinityfields,"IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 1, pp. 172-186,2021.
[3] H.-S.Fang,J.Li,H.Tang,C.Xu,H.Zhu,Y.Xiu,Y.-L.Li,andC.Lu,"AlphaPose: Whole-body regional multi-person pose estimation andtrackinginrealtime,"IEEETrans.PatternAnal.Mach.Intell.,vol.45,no. 6, pp. 7157-7173, 2023.
[4] J. Shin, A. S. M. Miah, R. Egawa, K. Hirooka, M. A. M. Hasan, Y.Tomioka, and Y. S. Hwang, "Fall recognition using a three-streamspatio-temporalGCNmodelwithadaptivefeatureaggregation,"Sci.Rep.,2025.
[5] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolutionrepresentationlearningforhumanposeestimation,"inProc.IEEECVPR, 2019, pp. 5693-5703.
[6] B. Jo and S. Kim, "Comparative analysis of OpenPose, PoseNet, andMoveNetmodelsforposeestimationinmobiledevices,"TraitementduSignal, vol. 39, no. 1, pp. 119-124, 2022.
[7] Google,"PoseestimationandclassificationonedgedeviceswithMoveNet and TensorFlow Lite," TensorFlow Blog, 2021.
[8] H.-S. Fang, S. Xie, Y.-W. Tai, and C. Lu, "RMPE: Regional multi-personposeestimation,"inProc.IEEEICCV,2017,pp.2334-2343.
[9] Z.Cao,T.Simon,S.-E.Wei,andY.Sheikh,"Realtimemulti-person2Dpose estimation using part affinity fields," in Proc. IEEE CVPR, 2017,pp.1302-1310.
[10] C.Zheng,W.Wu,C. Chen,T.Yang,S.Zhu,J.Shen,N.Kehtarnavaz,andM.Shah,"Deeplearning-basedhumanposeestimation:Asurvey,"ACM Comput. Surv., vol. 56, no. 1, 2023.
[11] E.Samkari,M.Arif,M.Alghamdi,andM.A.AlGhamdi,"Humanposeestimationusingdeeplearning:Asystematicliteraturereview,"Mach. Learn. Knowl. Extr., vol. 5, no. 4, pp. 1612-1659, 2023.
[12] Z. Huang, Y. Liu, Y. Fang, and B. K. P. Horn, "Video-based falldetectionforseniorswithhumanposeestimation,"inProc.4thInt.Conf. Universal Village (UV), 2018, pp. 1-4.
[13] Y. Wang and T. Deng, "Enhancing elderly care: Efficient and reliablereal-timefalldetectionalgorithm,"DIGITALHEALTH,vol.10,2024.
[14] T.Riahi,M.A.H.Bappy,andM.M.Islam,"ElderFallGuard:Real-timeIoTandcomputervision-basedfalldetectionsystemforelderlysafety," arXiv:2505.11845, 2025.
[15] E.R.Sykes,"Next-generationfalldetection:Harnessinghumanposeestimation and transformer technology," Health Systems, 2024.
[16] V.MaliandS.Jaiswal,"Pose-basedfalldetectionsystem:Efficientmonitoring on standard CPUs," arXiv:2503.19501, 2025.
[17] H.ZhengandY.Liu,"LightweightfalldetectionalgorithmbasedonAlphaPoseoptimizationmodelandST-GCN,"Math.Probl.Eng.,vol.2022,2022.

[18] V. Kong, S. Soeng, M. Thon, W.-S. Cho, A. Nayyar, and T.-K. Kim,"PIFR: A novel approach for analyzing pose angle-based humanactivitytoautomatefalldetectioninvideos,"PLoSOne,vol.20,no.6,2025.
[19] E.Alam,"GMDCSA-24:Adatasetforhumanfalldetectioninvideos,"2024.
[20] T.S.Kavya,"Visionbasedreal-timemonitoringsystemforelderlyfall detection," Comput. Syst. Sci. Eng., vol. 42, no. 1, 2023.
[21] P.Jesudhas,"Computationallyefficientnon-intrusivepre-impactfalldetection system," arXiv:2507.03705, 2025.
[22] C.S.T.Hii,K.B.Gan,N.Zainal,N.MohamedIbrahim,S.Azmin,S.
[23] H.MatDesa,B.vandeWarrenburg,andH.W.You,"Automatedgaitanalysisbasedonamarker-freeposeestimationmodel,"Sensors,vol.23, no. 14, 2023.
[24] V.Wagh,M.W.Scott,J.W.Andrushko,C.B.Jones,B.C. Larssen, L. A. Boyd, and S. N. Kraeutner, "Using MediaPipe to track upper-limbreachingmovementsafterstroke:Aproof-of-principlestudy,"J.NeuroEng. Rehabil., vol. 22, no. 1, 2025.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)