1 LITERATURE REVIEW

This literature review explores the latest developments in mmWave radar-based HAR systems. It encompasses of the background of both mmwave radar and human action recognition, delving into their applications and challenges. Also key datasets used, essential data processing techniques, analyze cutting-edge model architectures, and explore practical applications. Through this comprehensive analysis, the aim is to provide insights into both current capabilities and future possibilities in radar-based HAR.

1.1 Introduction to Human Action Recognition

HAR is a critical technology that has seen substantial change over the past few years. It encompasses of various methodologies and technologies aimed at automatically identifying and classifying himan activitues through various sensing systems. HAR has emerged as a game-changing technology that helps us understand and analyze human activities in real-time. With the usage od mmwave radar, these systems have become more robust, ensuring privacy and accuracy. **Kong2022** emphasized on how HAR has evolved from simple motion detectuon to more sophisticated activity classification platforms, incoporationg advanced artificial intelligence and sensor fusion technologies. Additionally, **Morshed2023** highlights that this evolution hgas been driven by a continous push towards more accurate, privacy preserving, and environmentally resillent solutions.

1.1.1 Evolution of HAR Technologies

As stated earlier from **Kong2022** HAR has significantly transformed over the past few years. The paper provides a comprehensive analysis of the evolution , from the traditional vision based approaches to more advances sensing technologies. This field also experienced growth with the integration of artificial intelligence and advances sensing modalities in modern har systems **Singh2021** . **Ariza2022** in their extensive review highlight how early HAR efforts relied on wearable sensors and basic data classification techniques, focusing on real-time monitoring for assisted living and healthcare applications. Additionally, in the earlt 2000s initial methods for HAR relied heavily on hand crafted features and simple statistical models **Aggarwal2011**. These include techniques such as template matching and motion analysis that were used to basic actions like walking and waving. These techniques struggled with background clutter and occlusion **Aggarwal2011**. This further transitioned to more sophisticated feature extraction techniques like Histogram of Oriented Gradients (HOG) and Motion History Images (MHI), where researches utilized machine learning classifiers such as Support Vector Machines (SVMs), to improve recognition accuracy in mid 2000s **Wang2020**. With the emergence of deep learning in 2010s, Convolutional Neural Networks (CNNs) began to dominate by enabling automatic feature extraction fdrom raw video data, this shift allowed for more complex actions to be recognised and improve robustness against various lighting and background **Karim2022 Ariza2022**. Recent developments focus on the integration of multiple modalities such as depth inform, skeletal data amd audio signals. The field is currently exploring real-time HAR applications, unsupervised and supervised learning and the use of generative adversarial networks (GANs) for data sugemtation **Aggarwal2011**.

1.1.2 Industry Applications and Use Cases

HAR technologies have diverse applications across multiple industries with unique requirements and challenges. In healthcare, the technology is instriumental for patient monitoring, fall detection and rehabiliation tracking**Appleby2007**. This can improve patient safety and care highlighting the importance of real-time monitoring response capabilites **Wang2019**. In consumer health technology, HAR is used for activity tracking, including step counting and sleep analysis, through smart wearables like fitness trackers and smartwatches, demonstrating its scalability for individual application **Niemann2022**. In addition to improving health, HAR is promoting accessibility through its use in sign language recognition, which helps the hearing impaired communicate by interpreting hand gestures**Niemann2022**. However, in industrial settings, context-awareness combined with

HAR greatly improves accuracy since systems may identify operations based on the position, identity, or environment of relevant items **Niemann2022** . The development of digital twins to mimic human behavior is made possible by this context-aware method in conjunction with cyber-physical systems, which expedites the creation of datasets for intelligent systems and expands HAR's industrial potential**Niemann2022**. Furthermore, it enhances efficiency in material handlying systems in automating recognition of complex human activites crucial to optimiing workflows in warehouses and distribution centers **Reining2019**. Also the use of motion capture systems for complexity of movements in industries warrants the use of human action recognition. **Ickstadt2022**.

1.1.3 Comparison with Vision-Based HAR

Computer vision has historically dominated HAR through camera-based methods, capturing high-resolution data for recognizing human activities in various domains, such as surveillance and healthcare. However, these systems have significant limitations, including reliance on ambient lighting, vulnerability to occlusions, and privacy concerns in sensitive environments.

In contrast, mmWave radar offers privacy-preserving capabilities by capturing only motion data without identifiable imagery, making it suitable for continuous monitoring in private spaces. Its ability to penetrate obstacles and perform accurately in low-light conditions makes it particularly advantageous in settings where computer vision falls short. This comparison underscores mmWave radar's unique position as a robust, privacy-friendly solution in HAR, especially when combined with advanced machine learning techniques.

1.1.4 Challenges in Traditional HAR Systems

Aggarwal2011 Ariza2022 Wang2020 Roshan2022 Raza2021 outlines primary challenges of traditional systems such as environmental variations, requiring controlled conditions for optiomal performance in real life scenariois. Primarily, these systems struggle with variability in viewpoints actions observed from different angles or perspectives often yield inconsistent recognition, making them unreliable in multi-camera or uncontrolled settings **Raza2021**. Complexity is further increased by inter-class similarity and intra-class diversity. Different people carry out the same activity in different ways, and some activities have characteristics that overlap, making it challenging for conventional systems to tell them apart. Their accuracy is restricted by this lack of flexibility, particularly in cases when feature sets are inflexible **Roshan2022**. Environmental factors are challenges as well. Since these methods are not robust enough to isolate human activity in the presence of complex or dynamic backgrounds, background instability and lighting conditions frequently cause problems with recognition. Action visibility is further altered by poor lighting, which also reduces the system's dependability in dimly lit or unevenly illuminated situations**Yu2020**. Finally, computational demands remain a barrier. Traditional methods often require high processing power, which is impractical for real-time applications. This trade-off between speed and accuracy limits their usability in live scenarios where immediate responses are critical. Overall, these limitations underline the need for more advanced, adaptable HAR systems to meet real-world demands**Aggarwal2011**.

1.2 Millimeter-Wave Radar Technology

1.2.1 Background of mmWave Radar

Millimeter-wave (mmWave) radar is a specialized radar sensor that radiates continuous transmission power, similar to a simple continuous-wave radar, by analyzing reflected signals **Liu2020**. mmWave radar systems operate in the frequency range of 30 GHz to 300 GHz and offer unique advantages, including high spatial resolution, material penetration capabilities, and robust performance across various environmental conditions **Appleby2007**; **AlHourani2018**. These characteristics make the technology suitable for HAR applications, as it can capture subtle human movements while maintaining privacy and operating effectively in different lighting conditions **Gu2023**.

The foundation of mmWave radar is the Frequency Modulated Continuous Wave (FMCW) radar, which measures the range, velocity, and angle of objects **Iovescu2020**. FMCW radar operates by transmitting a frequency-modulated signal, known as a chirp, and analyzing the reflected signals to determine target parameters **Liu2020**. The fundamental principle involves transmitting a linear frequency sweep starting at a carrier frequency **fc,** with bandwidth **B** over duration **Tc**.

Range estimation in FMCW radar is achieved through the analysis of the Intermediate Frequency (IF) signal, generated by mixing the transmitted and received signals **Liu2020**. The distance to a target can be calculated using the equation:

$$d = \frac{f_{IF} c}{2S} \tag{1}$$

where ***c*** is the speed of light, ***fIF*** is the ***IF*** signal frequency, and ***S*** represents the frequency slope ***(B/Tc)*** **Wang2020**.

1.2.2 Signal Processing in mmWave Radar

mmWave radar systems employ sophisticated signal processing techniques to enhance performance, including **Feng2023**:

1. **Range Fourier Transform (Range-FFT):** The Range-FFT is a fundamental signal processing technique in mmWave radar systems **Wang2020**. This transformation converts the time-domain IF signals into the frequency domain, enabling accurate target range detection by analyzing reflections **Iovescu2020**. Studies have shown that the Range-FFT's effectiveness in HAR applications stems from its ability to precisely locate human subjects in the radar's field of view **Huang2022**. The transformation displays discrete peaks in the frequency spectrum corresponding to targets at specific ranges:

   $$d = fIF/2S$$

   where ***d*** represents the target distance, ***fIF*** is the ***IF*** signal frequency, and ***S*** denotes the chirp slope. Research demonstrates that this technique can achieve centimeter-level accuracy in human position detection, making it valuable for fine-grained motion analysis in HAR **Lee2023**.

2. **Doppler Fourier Transform (Doppler-FFT):** Doppler-FFT is critical in mmWave-based HAR systems for capturing human motion dynamics **Ige2022**. This method estimates target velocity by analyzing phase variations between successive chirps **Iovescu2020**. Studies show that Doppler-FFT provides unique motion signatures for different human activities; for example, walking patterns generate distinctive Doppler spectrograms, significantly different from falling motions, thus enabling reliable fall detection **Feng2023**. The velocity resolution can be as precise as 0.25 m/s, distinguishing subtle movement variations **Gu2023**. Doppler-FFT's capabilities have enabled algorithms to recognize diverse activities accurately while preserving privacy **Lee2023**.

3. **Angle of Arrival (AoA) Estimation:** AoA estimation uses multiple receiver antennas to determine the target's angular position, essential for tracking human movements in 3D space. This technique is particularly useful in applications requiring spatial awareness, such as multi-person tracking in public areas or monitoring within confined spaces.

Suggested Addition: Point cloud data processing plays a critical role in mmWave HAR applications, where 3D point clouds generated from radar returns are processed using clustering and noise reduction methods. These techniques help isolate relevant human movement data from background noise, enhancing HAR accuracy and enabling more effective real-time activity recognition.

1.2.3 TI AWR1843 Development Platform

The Texas Instruments AWR1843 mmWave radar sensor has gained significant traction in HAR due to its high-resolution motion data capture capabilities in a non-invasive manner **Ti.com2021**. The AWR1843 is a standalone FMCW radar sensor, capable of implementing a 3TX, 4RX system with integrated PLL and ADC converters using TI's low-power 45nm RFCMOS technology **Ti.com2021**. The AWR1843 achieves range resolution down to 4 cm and velocity resolution up to 0.25 m/s, critical for distinguishing subtle human actions across various environmental conditions **Huang2022**.

Recent advancements in machine learning models, such as multi-head deep attention-based networks, have enhanced the accuracy and efficiency of HAR systems using mmWave radar. For instance, Raza et al. (2021) developed a multi-head deep attention-based long short-term memory (LSTM) neural network that achieves high HAR accuracy by focusing on critical temporal features **Raza2021**. Such integration of deep learning models with the AWR1843 supports accurate, non-invasive, and privacy-preserving HAR, particularly in healthcare and personal monitoring **Morshed2023**.

1.2.4 Evolution of mmWave Radar for HAR

Historically, radar systems were used in military and large-scale industrial applications, largely due to their size and complexity. However, advancements in semiconductor technology have led to compact, high-resolution mmWave radars for consumer-level applications **AlHourani2018**.

The classification of mmWave radars into Long-Range Radar (LRR) and Short-Range Radar (SRR) illustrates their versatility. LRR, with detection ranges exceeding 200 meters, is ideal for applications like vehicle anti-collision systems, while SRR is effective in short-range applications, offering reliable data even in adverse weather conditions **Iovescu2020**. This shift from industrial to consumer applications has unlocked new research opportunities in HAR, especially in environments where privacy is a concern (Wang et al., 2020).

1.2.5 mmWave Radar's Role and Application in HAR

The unique capabilities of mmWave radar make it superior to other sensor modalities across various scenarios. While camera-based systems provide high-resolution imagery, they depend on ambient light and often fail in low-light conditions. In contrast, mmWave radar operates in total darkness and penetrates obscurants such as smoke or fog, making it invaluable for HAR in diverse environments **Iovescu2020**.

Additionally, mmWave radar preserves privacy by offering non-intrusive monitoring without the need for physical contact or visual capture, making it ideal for monitoring in sensitive settings, such as homes and healthcare facilities **Gu2023**. When compared directly, mmWave radar outperforms traditional camera systems in multi-person tracking scenarios, with radar signals penetrating through objects, providing consistent recognition even in obstructed environments **Chen2021**.

1.2.6 Challenges in mmWave HAR

Despite its advantages, mmWave radar technology faces challenges in HAR. High-resolution data processing is computationally demanding, posing limitations for real-time applications in resource-constrained settings. Differentiating between similar human activities, such as walking and running, can also be difficult, especially in crowded scenarios where multiple subjects are present.

Furthermore, interference from other devices operating in the same frequency band remains a challenge. As more devices adopt mmWave technology, interference mitigation will require advanced signal processing to maintain reliable performance.

Current research focuses on addressing these challenges by enhancing spatial and temporal resolution, optimizing signal processing algorithms for real-time performance, integrating radar data with machine learning, and reducing power consumption.

1.3 Advanced Processing Techniques

Effective data processing is fundamental HAR as it enables the extraction, filtering, and interpretation of relevant information from noisy sensor data, which is especially crucial when working with mmWave radar signals. In this project, advanced clustering, filtering, and noise reduction techniques were chosen to enhance accuracy and robustness, addressing the challenges inherent to mmWave radar data processing. Specifically, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Kalman filtering with Hungarian algorithim for data association were selected for their distinct capabilities in noise handling, object tracking, and efficient data association. These methods support a streamlined and accurate HAR system capable of tracking multiple individuals in real-time.

1.3.1 Density-Based Clustering (DBSCAN)

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a clustering algorithm particularly suited for identifying dense regions of data points while classifying low-density points as noise. This feature is beneficial in HAR applications that rely on radar data, where raw sensor data is often cluttered with noise or irrelevant signals. DBSCAN operates by finding clusters based on density, allowing it to discover clusters of arbitrary shapes, which is particularly useful when detecting non-linear human movement patterns **Raza2021**. Unlike traditional clustering algorithms such as K-means, DBSCAN does not require the number of clusters to be specified in advance, making it adaptable to dynamic environments where the number of detected objects can vary. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a fundamental clustering algorithm that excels in discovering clusters of arbitrary shapes while effectively handling noise in spatial databases (Deng, 2020). The algorithm's core principle revolves around two key parameters: Eps (radius) and MinPts (minimum points threshold). It classifies points into three categories: core points, border points, and noise points, based on the density of their neighborhood (Zhao et al., 2019).

A point is considered a core point if its neighborhood within radius Eps contains at least MinPts points. This density-based approach makes DBSCAN particularly suitable for point cloud processing as it can effectively handle varying cluster shapes and densities commonly encountered in radar data (Deng, 2020). The algorithm's ability to automatically identify and filter out noise points is especially valuable in mmWave radar applications, where signal reflections and environmental factors can introduce significant noise.

In the context of mmWave radar point cloud processing, DBSCAN offers several advantages. The algorithm's density-based approach aligns well with the characteristics of radar-generated point clouds, where human subjects typically appear as dense clusters of points against a sparser background (Zhao et al., 2019). This makes it effective for human detection and tracking applications, as demonstrated by Zhao et al. (2019) who achieved reliable multiple-person tracking using DBSCAN-based clustering.

When applied to mmWave radar data, DBSCAN helps in: 1. Distinguishing between actual human subjects and noise in the point cloud 2. Handling multiple subjects simultaneously through effective cluster separation 3. Maintaining tracking consistency through density-based cluster association

However, the implementation requires careful parameter tuning. As noted by Zhao et al. (2019), points corresponding to human subjects are typically coherent in the horizontal (x-y) plane but more scattered along the vertical (z) axis. This observation led to the modification of the traditional Euclidean distance metric used in DBSCAN to better suit mmWave radar point clouds:

‘‘‘ D(pi, pj) = (pi$_x - pj_x)^2 + (pi_y - pj_y)^2 + *(pi_z - pj_z)^2$‘‘‘

where is a weighting parameter that reduces the contribution of vertical distance in the clustering process (Zhao et al., 2019). This modification significantly improves the algorithm's performance in human tracking applications.

The effectiveness of DBSCAN in mmWave radar applications has been demonstrated through various metrics. For instance, Zhao et al. (2019) achieved median tracking accuracy of 0.16m and identification accuracy of 89

Zhao2019 Deng2020 Ahmed2016 In the context of this project, DBSCAN was selected to initially segment mmWave radar data into clusters that represent distinct objects or individuals, while ignoring noise. The algorithm's robustness against noise and ability to handle arbitrary-shaped clusters make it ideal for separating human activity data from background clutter, which is essential in real-world HAR scenarios **Huang2022**. As previous research has shown, DBSCAN's performance in detecting activity clusters in high-noise environments significantly improves the accuracy of HAR systems by isolating relevant data points from extraneous signals, thus enhancing the reliability of subsequent stages in the recognition pipeline **Morshed2023**.

It involves the use of two parameters; - minPts : also known as the minimun sample, is the number of points clustered together for a fegioin to be considered dense - Epilson (eps e); the istance measured usedd to locate the points in the neighborhood of any point

1.3.2 Kalman Filtering for Smoothing and Prediction

Kalman filtering is a powerful recursive algorithm used to estimate the state of a dynamic system from noisy measurements. In HAR applications, Kalman filtering is typically employed to smooth sensor data and predict the future positions of detected objects based on their current trajectory. The Kalman filter operates by continuously updating its estimates based on new incoming measurements, minimizing the variance of the estimate through a series of prediction and correction steps **Gu2023**. This filtering technique is especially useful in scenarios where precise motion tracking is required, such as distinguishing small human actions or movements.

In this project, the Kalman filter was integrated to enhance the accuracy of position and velocity estimates derived from radar data, ensuring that movement trajectories are smooth and realistic. The application of Kalman filtering addresses the inherent noise in mmWave radar data, which can result in jittery or inaccurate motion paths if left untreated. Furthermore, as human activities often involve complex and rapid motions, Kalman filtering's predictive capabilities allow the system to track movements more effectively, maintaining accuracy even when individuals are moving at varying speeds or abruptly changing directions **Zhao2019**. This predictive element is crucial for real-time HAR, where tracking consistency and accuracy are paramount.

Data Association with the Hungarian Method The Hungarian method is an algorithm widely used for solving assignment problems, particularly effective in associating detected objects across consecutive frames. In HAR applications, the Hungarian method is valuable for tracking individuals over time by associating new detections with previously established trajectories, minimizing instances of identity switching or loss of tracking. The method's efficiency in solving matching problems makes it a strong candidate for use in conjunction with Kalman filtering, which provides predictions for object positions in the next frame **Karim2024**.

In this project, the Hungarian method was chosen to maintain continuity in tracking multiple objects detected by the radar, especially in environments with several moving targets. By pairing this method with Kalman filtering, the system can accurately link detections across frames, even when movement patterns are erratic or when individuals temporarily occlude each other. This approach mitigates one of the common issues in multi-target tracking: the ambiguity in data association, which can arise in scenarios with complex human interactions or overlapping movements. Utilizing the Hungarian method ensures that each detected individual maintains a unique identifier over time, which is critical for generating consistent and interpretable activity data **Gies2022**.

1.3.3 Integration and Justification of Selected Techniques

The combined use of DBSCAN, Kalman filtering, and the Hungarian method in this project addresses the primary data processing challenges encountered in HAR using mmWave radar: clustering relevant data points, reducing noise, and ensuring accurate object tracking over time. DBSCAN's ability to isolate dense clusters amidst noise makes it ideal for initial object detection in radar data, especially in complex or noisy environments where background clutter can interfere with accurate detection. Kalman filtering then enhances the system by smoothing and predicting the movement of each detected object, providing more accurate real-time tracking of human activities **Morshed2023**. Finally, the Hungarian method facilitates consistent multi-object tracking across frames, maintaining the integrity of activity sequences by effectively associating detections with ongoing trajectories.

This approach leverages each method's strengths to address the unique challenges posed by HAR in real-world settings. The robust clustering from DBSCAN, the predictive accuracy of Kalman filtering, and the efficient data association by the Hungarian method together create a comprehensive data processing framework. This framework not only improves the clarity and reliability of radar data but also enables accurate HAR that respects privacy by not requiring visual data capture **Yang2020**; **Kong2022**. Consequently, the system achieves a high level of accuracy in capturing and interpreting complex human activities, which is essential for both research and practical applications in fields such as healthcare monitoring, security, and smart environments.