**Simulating Intelligent Elderly Care Systems**

Nate Bachmeier

CMP-9701: Precandiacy Prospectus for Computer Science

December 26, 2021

Northcentral University

# Table of Contents

[Table of Contents 2](#_Toc87544098)

[Background 4](#_Toc87544099)

[Problem Statement 5](#_Toc87544100)

[Purpose Statement 6](#_Toc87544101)

[Research Questions 8](#_Toc87544102)

[Hypotheses 9](#_Toc87544103)

[Research Methodology 9](#_Toc87544104)

[Artifacts 10](#_Toc87544105)

[Data Generation 10](#_Toc87544106)

[Environmental Simulation 10](#_Toc87544107)

[Extracting Intents 11](#_Toc87544108)

[Feedback Mechanism 11](#_Toc87544109)

[Artifact Requirements 11](#_Toc87544110)

[~~User Roles~~ 12](#_Toc87544111)

[System Architecture 12](#_Toc87544112)

[System Reliability 13](#_Toc87544113)

[Contributions 13](#_Toc87544114)

[Measurements and Evaluation 14](#_Toc87544115)

[Data Collection Process 14](#_Toc87544116)

[Evaluation Process 15](#_Toc87544117)

[Benchmarking 15](#_Toc87544118)

[Summary 16](#_Toc87544119)

[References 18](#_Toc87544120)

[Annotated Bibliography 22](#_Toc87544121)

[Human Activity Recognition 22](#_Toc87544122)

[Integrating IoT Systems 23](#_Toc87544123)

[Enhancing Security 23](#_Toc87544124)

[Healthcare and Cloud 24](#_Toc87544125)

**Simulating Intelligent Elderly Care Systems**

This constructive research project seeks to raise the elderly care quality bar while minimizing costs. Using Human Activity Recognition (HAR) through computer vision and machine learning, the solution tracks the patient’s skeletal structure. After observing behaviors and determining intents, the system orchestrates Cyber-Physical Systems (CPS) to provide further assistance.

# Background

A demographic timebomb will create significant pressure on the global health care system because people live longer, have fewer children, and medical costs continue to increase (Piggott, 2016; Stone, 2017). When patients cannot afford the required care, the quality decreases, or social programs must fund the difference. Demographic specialists predict that by 2050 nearly “80% of the global elderly population will be from low- to middle-income countries (Muhsin, Munyogwa, Kibusi, & Seif, 2020, p. 1).” Economic constraints within those countries will limit the effectiveness of their welfare programs and the availability of adequate services. Additionally, over one billion globally have a limiting disability that requires additional support (Morris, 2008). Medical facilities need mechanisms to defuse the situation by reducing costs and deferring the transition to assisted living centers.

Inversely, the explosive growth across IoT, Cloud, Big Data, and Mobile (ICBM) continuously decreases costs and enables new scenarios. These technologies will revolutionize the health care and wellbeing industries. Academic and commercial vendors are continuously delivering innovations across these domains. However, mainstream offerings primarily focus on measuring simple body metrics (Koreshoff, Robertson, & Leong, 2013). While these products provide incremental value, they do not move the needle. ~~Nearly eight years later, the industry myopically drives toward wearable IoT devices (Tun et al., 2021). Researchers concentrating on these areas make sense due to the low barrier to entry. Though, that same ease is commoditizing the products selection and stifling creativity.~~

Technology within special needs and elderly care settings has unique challenges and requirements (Ferati, Kurti, Vogel, & Raufi, 2016). For example, these solutions often support safety functions and require uninterrupted monitoring. ~~Wearable IoT has inherent risks that the patient will disable or forget the device. This situation is particularly profound in patients with memory impairments like dementia (Wilson, 2017). Additionally, the obtrusive nature of wearable technologies makes them impractical for extended duration scenarios (Razzaq et al., 2020; Singla et al., 2010).~~ Alternatively, specific vendors utilize voice-enabled Personal Digital Assistants (PDA) (e.g., Amazon Alexa). These products effectively set reminders and record activities (Tan et al., 2020). However, several situations cannot leverage voice technologies, such as non-native speakers and individuals with vocal disorders.

Assisted living facilities use trained nurses to mitigate these issues. Having a human inspect the patient visually is an effective but expensive tool. The median compensation rate for registered nurses is $75,330 annually ($36.22 per hour) (US Bureau of Labor Statistics, 2020). Due to the high cost, few patients have private nurses and receive fractional supervision. Instead, patients could receive continuous observation at lower costs using Computer Vision. Artificial Intelligence and Machine Learning (AI/ML) models can observe patient behavior and react accordingly.

# Problem Statement

Senior citizens live longer than ever before and want to defer moving into nursing homes until later in life. Transitioning into elderly care comes as a double edge sword. On the one hand, nurses can provide 24-hour supervision. This assistance could mean the difference between life and death (e.g., during a fall). However, on the other hand, the medical services are prohibitively expensive, nearing $90,000 annually (Tan et al., 2020). Additionally, these medical facilities lack the personalization available within one’s home. Further, specific individuals with diseases like dementia and Alzheimer’s require even greater levels of attention.

Traditionally, addressing the situation requires increasing human capital, such as more traveling nurses or family member oversight (Westergren et al., 2021). However, this solution increases health care costs and collects limited patient health data. In addition, these infrequent visits might miss critical issues, especially with those most reluctant to relocate.

Modern solutions must bridge the differentiation between remaining in the home and still receiving the attentiveness typically found in assisted living facilities (Tan et al., 2020). When this gap narrows, it enables the patient to remain in familiar settings for more prolonged periods. That situation has numerous psychological benefits (e.g., higher morale) and economic impacts (e.g., deferring private health care costs). For instance, patients with memory impairment might forget to empty the dishwasher, take medication, or bathe regularly. Medical facilities can address these challenges through real-time video monitoring services that analyze patients’ actions and recommend care. After detecting an issue, the system alerts the person using Text-to-Speech (TTS) services (e.g., Amazon Alexa) and other Cyber-Physical Systems (CPS).

# Purpose Statement

This constructive design research project defines and implements an Elderly Care Smarthome Operating System (ECSOS). The ECSOS will provide central core services for bringing world-class assisting living care into a resident’s home, such as identity management, patient action tracking, consistent Cyber-Physical control plane, and privacy functions. While this specific project examines elderly care, the implications are generalizable to other scenarios. Those scenarios encompass childcare (e.g., babysitting), school safety systems, and virtual office secretary situations, to name a few.

Building these capabilities and verifying them at scale is challenging. First, the research must find patients willing to share continuous in-home video streams. In addition to the privacy concerns, it would be difficult for others to reproduce the findings. Third, validating the solution against numerous home layouts requires significant effort. Lastly, purchasing and configuring hardware components is prohibitively expensive in terms of time and money (Elloumi et al., 2020; Das et al., 2019). This project mitigates these issues through a virtual world simulation process.

The Robot Operating System (ROS) is a framework for writing robot software (Stanford Artificial Intelligence Laboratory et al., 2018). It exposes features for rapidly designing complex cyber-physical interactions through a Message Passing Interface (MPI). The meta-operating system integrates with physics engines (e.g., Gazebo) and machine learning platforms (e.g., OpenAI Gym). Additionally, developers can package these tools into containerized workloads and leverage public cloud services (e.g., Amazon Web Services and Microsoft Azure). The cloud enables researchers to validate their designs in numerous world permutations efficiently and economically. Furthermore, the software can test situations that are not practical or feasible within the physical world (e.g., set the kitchen ablaze). Together, these different technologies culminate into an elegant solution that monitors, predicts, and responds in real-time to patient needs.

This dissertation leverages these tools to implement an intelligent home simulation environment. Next, it will populate the virtual home with ROS devices and sensors to observe and respond to ROS actors (patients). The actors will perform animation sequences based on motion-capture records. Lastly, the researcher will assess the observations and responses against the labeled data. While this specific test scenario focuses on elderly care, the solution is broadly applicable to any Cyber-Physical simulation.

# Research Questions

Researchers are innovating across health care using Internet of Things (IoT) devices (Tun et al., 2021). ~~Their efforts predominately focus on wearable technologies that attach sensors to the patient. Wearable technologies face significant competition because these solutions have a low barrier to entry, economical pricing, and are mass-producible.~~ ~~However, these products lack elegance due to restricting movement and necessitating the patient always to carry the device.~~ Additionally, the saturated market causes each iteration to produce less incremental value.

In contrast, high-quality research must be challenging, elegant, and move the needle (Zeller, 2014). Meeting these requirements necessitates a different approach, such as utilizing cameras and real-time video processing to deliver a superior experience. However, video-centric systems encounter more complexity in several aspects. For instance, patients can freely move around their residence and change its configuration (e.g., move furniture or turn off a light). Addressing the noise within these dynamic environments is challenging and creates multiple research questions.

**R1** – How are researchers minimizing noise in their video streams? An efficient process must exist to analyze short videos and extract the subject’s *intent*. This mechanism must reliably handle noisy data (e.g., out-of-focus images) and variable input (e.g., distance to the camera).

**R2** – How can the extracted intents best *interface* with Cyber-Physical Systems (CPS)? Nurses at assisted living centers provide a helping hand literally and figuratively. Smart devices must serve this same function across a range of tasks (e.g., medication management).

**R3** – How can those interfaces ensure patient *confidentiality*? Patients will only use a continuous video recording solution if they trust its security and privacy controls. There must be explicit and deliberate decisions regarding how information is stored and replicated.

**R4** – How can central administrative teams most efficiently *scale* across global and domestic sales territories? Healthcare workers can remotely deliver world-class services because the homes contain CPS systems for routine tasks (e.g., monitoring patient falling). Competitive businesses can leverage this capability to decrease costs, increase profit margins, and maintain quality standards.

# Hypotheses

There are existing mechanisms to address each system requirement. However, those capabilities exist as isolated components. An Elderly Care Smarthome Operating System (ECSOS) can create a consistent control plane and test framework that brings these services together. The ideal driver of such a solution is real-time video processing at the edge. This approach promotes security and privacy by minimizing external data transfers. ~~Additionally, small appliances built on the Raspberry PI platform have sufficient resources to drive computer vision models.~~

# Research Methodology

Before starting any significant undertaking, there needs to be a formal project plan that scopes the intent. Additionally, the plan must define mechanisms to measure the success and impact of those efforts. It can be challenging or impossible to prove that the project’s outcomes are efficient without those prerequisites.

## Artifacts

Artifacts are a principal component of the constructive design methodology. This research proposal includes four core subsystems. Together, these deliverables generate data, simulate interactions, make predictions, and score outcomes.

### Data Generation

ROS actors represent the patients within the simulation environment, which performs an animation sequence while moving around the house. These animations originate from Microsoft Kinect motion-capture videos and map to a hierarchial action-space taxonomy. The action-space describes specific behaviors (e.g., walking versus sitting) and any derived actions (e.g., sitting on a chair versus couch). There are virtually infinite sequences, making it challenging to record the entire universe of movement. Instead, a randomization process initializes from a recording and mutates model-joint characteristics such as flexibility, strength, and weight. This approach both increases taxonomy coverage and prevents overfitting the limited data.

### Environmental Simulation

ROS worlds represent the patient’s home or apartment and define models’ placement (e.g., actors and furniture), actor configuration, and devices (Bipin, 2018). Researchers use physics simulators (e.g., Gazebo) to examine interactions between these various components. For instance, the actor might perform walking to the kitchen table. Each camera will capture frames from its vantage point and transmit them to a message bus during this sequence. Next, AI services subscribe to the event stream and process the visual data. Suppose the service detects a valuable signal (e.g., the refrigerator door is left open). In that case, it can post a notification to another message bus to mitigate the situation (e.g., use voice assistant).

Validating these interactions requires an ability to reconfigure these worlds without significant effort. World templating tools (e.g., AWS RoboMaker) can dynamically generate environments that meet a specification (AWS, 2021). This capability allows the researchers to focus on creating custom sensors and algorithms, not positioning furniture. That also means this dissertation should have a stronger emphasis on ROS components and world templates, not reinventing standard tooling. These components must implement an asynchronous and loosely coupled architecture.

### Extracting Intents

A machine learning algorithm will classify and annotate the video’s contents. There are several potential implementations (e.g., Open Pose versus Das et al.’s approach). The AI model(s) must perform within the hardware constraints of an edge appliance. For instance, the simulated home might parallel process dozens of cameras and sensors. Otherwise, the system will require remote compute configurations (e.g., public cloud), which raises security and privacy concerns.

### Feedback Mechanism

Third, the system records all observations and responses for offline comparison against the labeled data. These records qualitative tags and aggregates centrally for offline analysis. The research team will use this database to assess the system’s accuracy and identify potential quality gaps. For instance, the responses could indicate that specific intent predictions are unreliable (e.g., the patient has a fork versus pen).

## Artifact Requirements

Customers will only use artifacts that meet their specific needs, are well designed, and are reliable. Researchers must meet these expectations with concise deliverable requirements.

### ~~User Roles~~

~~Aside from patients, there are three additional user roles: nurses and healthcare providers, family members, and administrators. These users can use a mobile app or web portal to access the relevant data. All operations from either UI (User Interface) require Authentication, Authorization, and Auditing (AAA). When systems mandate AAA enforcement, it prevents negligence or malicious actions while increasing transparency. It is also critical that the patient maintains control of their privacy. For example, they might want to share a weekly aggregate health report with family members, not verbose details.~~

### System Architecture

~~Elderly Care SOS requires cameras, network storage, and a custom-built appliance (see Figure 3). Optionally patients can extend the system with various CPS device integrations (e.g., remote smoke detectors). The appliance must have enough computing and storage resources to perform model predictions, persist state, and execute several micro-services. Periodically, the on-premise system needs to synchronize with an external cloud component. These synchronization operations include sending status reports, downloading updates, and issuing assistance requests.~~

Figure 3: Abstract Design

Diagram

Description automatically generated

### System Reliability

~~The architecture’s components communicate over WiFi, Zigbee, and Bluetooth protocols. Time-sensitive messages (e.g., the subject has fallen) require a primary and secondary communication channel, such as phone line or mobile phone pairing. These messages are likely to encounter transfer failures due to radio interference or devices being offline. There must be support within the message buses to cache and reattempt any delivery failures using exponential backoff policies. Otherwise, the state management’s perspective can become distorted.~~

~~The appliance must locally run several services that handle core scenarios like identity and message routing. Developers can also load custom extensions that subscribe to event notifications. Those various subsystems require isolation and controls to limit the blast radius of a specific failure. An industry-standard approach would be to use micro-service designs and container orchestration technologies (e.g., Kubernetes)(Wen et al., 2020). These products can manage fail-over replicas and promptly restart crashed instances.~~

## Contributions

The core contribution to the body of knowledge is ~~the case study using the proof-of-concept design~~. Existing publications review each component within ECSOS under distinctly different use cases (e.g., sports injuries). Das et al. (2019) explain that those resources are not directly reusable, and implementations must use domain-specific labeled content. This design requirement necessitates compositing a new solution from custom and open-source software.

Second, the research ~~produces a purpose-built machine learning algorithm for elderly care action recognition. This deliverable also includes quantitative metrics that describe the algorithm’s resource utilization and F-measure accuracy. Data scientists use F-measurements as a “way of combining the precision and recall of the model, and [defines] the harmonic mean of the model’s precision and recall~~ (Wood, n.d.)~~.” Researchers can make trade-offs in their solution to optimize this value for their specific scenario. For instance, a critical health management system might enforce higher penalties on false negatives than false positives.~~

# Measurements and Evaluation

There must be formal mechanisms to collect service telemetry, evaluate its accuracy, and compare against third-party benchmarks.

## Data Collection Process

Most information enters the system ~~through the WiFi cameras~~. ~~Ideally, those cameras are accessible only through a dedicated Virtual Local Area Network (VLAN). This recommendation protects the unencrypted Real-Time Streaming Protocol (RTSP) from eavesdropping and tampering attacks.~~ After the video clip is available, ECSOS must process it through several machine learning models (e.g., facial recognition, object detection, and action recognition). These metadata annotations persist into a time-series database. ~~Lastly, populating the database requires the patient(s) to behave normally and let the system collect the video recordings.~~

~~Initially, the system will not have any training data and cannot make recommendations. However, researchers can accelerate data labeling with online products like Amazon SageMaker Ground Truth~~ (Amazon, 2021)~~. This service offers clustering capabilities to group related artifacts and streamlines manual tasks. Alternatively, users can crowd-source labeling jobs through Amazon Mechanical Turk.~~

## Evaluation Process

There must be feedback loops that confirm that the predictions and recommendations are accurate. Without this capability, it would be challenging to discover issues and prioritize machine learning model changes. ~~One potential solution is to collect these responses through a patient mobile app. App users can enumerate previous recordings and see the associated metadata. Those filmings are subject to a retention policy that automatically deletes old content. If they disagree with the predictions, they can make corrections inline. After making the manual update, the user can help improve the experience by submitting the footage to the ECSOS Cloud.~~

The ECSOS Cloud maintains a history of all incorrect predictions. Data scientists can review those responses, look for patterns (e.g., mixing up two actions), and make the necessary modifications. ~~There must be mechanisms to include user feedback and avoid introducing biases into the model~~ (García-Pérez, 2012)~~. For example, an individual user could post thousands of feedback comments containing inaccurate data change requests. This situation could regress other users’ experiences~~. Similar biases can enter the system due to insufficient test subjects’ racial diversity (e.g., only validating white men).

## Benchmarking

Numerous Human Activity Recognition (HAR) benchmarks exist with varying frame rates, actions, actors, backgrounds, resolutions, and problem domains (Singh & Vishwakarma, 2018). However, most benchmarks also focus on high-intensity outdoor sports footage (Das et al., 2019). Since those behaviors are very different from low-intensity indoor movements, they are not directly usable. Instead, several publications choose to define movement taxonomies and curated lists of expected behaviors. Afterward, the project’s quality is proportional to its ability to cover those actions. Additionally, the solution should be extensible and support more actions in future versions.

# Summary

Senior citizens and other debilitating populations are living longer and want to defer transitioning into assisted living facilities. Delaying the move can save nearly ninety thousand dollars annually and provide additional value (e.g., personalized comfort). However, those savings typically require care degradation. ~~Researchers and designers are mitigating these issues through wearable technologies (e.g., smartwatches). These solutions are impractical for extended periods due to restricting movement, requiring carrying the device, remembering to charge it, among other limitations. Instead,~~ the state of the art solutions utilizes real-time video processing to provide an equal or better experience without being obtrusive.

Computer vision strategies typically employ a skeletal tracking process. This process must first convert the patient’s skeleton into 3D space then assess the changes. That operation requires a combination of CNN, RNN, and ensemble algorithms. Multiple challenges arise while implementing these pipelines. For instance, there are minimal open datasets for analyzing indoor low-activity behaviors. This limitation causes researchers to build everything from step one forward.

Luckily, there are existing tools and services to simplify the development experience when designing complex behaviors. After determining the patient’s *intent*, the operating system must respond with an appropriate response. This reaction frequently requires manipulating the physical world from the virtual domain. Cyber-Physical Systems create those communication bridges, enabling data analytics to manipulate the physical realm. There are several well-established patterns for approaching these challenges, ~~but the ecosystem is fractured (e.g., requires multiple vendors).~~ Researchers must minimize adding complexity within this complicated problem space.

Lastly, unlocking these challenges could enable an extremely competitive service offering. Today, assisted living centers must employ numerous staff members and maintain the physical buildings. Instead, the ECSOS enables centralized nursing teams to scale across sales territories while reducing costs and increasing efficiencies. Additionally, since the patients remain within their residence, the business can leverage smaller facilities. Furthermore, the service is marketable to a broader audience. While traditional living centers only care for end-of-life patients. ECSOS applies to any widow or elderly parent. Children are often hesitant to “put their mom in a home” but will gladly subscribe to the piece of mind that “someone is watching mom.” The business can tap into those emotions to produce win-win scenarios.

# References

Aguida, M., Ouchani, S., & Benmalek, M. (2020). A review on cyber-physical systems. *Enabling Technologies: Infrastructure for Collaborative Enterprise* (pp. 275-278). Virtual: IEEE. doi:10.1109/WETICE49692.2020.00060

Amazon. (2021, June 11). *Amazon Go*. Retrieved from Amazon: https://www.amazon.com/b?node=16008589011

Amazon. (2021). *Use Aamzon SageMaker Grouth Truth to Label Data*. Retrieved from Amazon: https://docs.aws.amazon.com/sagemaker/latest/dg/sms.html

Amin, S., Salahuddin, T., & Bouras, A. (2020). Cyber-Physical Systems and Smart Homes in Healthcare: Current State and Challenges. *International Conference on Informatics, IoT, and Enabling Technologies* (pp. 302-309). Virtual: IEEE. doi:10.1109/ICIoT48696.2020.9089638

AWS. (2021). *AWS RoboMaker*. Retrieved from Amazon Web Services: https://aws.amazon.com/robomaker/

Bell, Koren, & Volinsky. (2009). *The BellKor solution to the Netflix Prize.* Retrieved from Netflix Prize: https://netflixprize.com/assets/GrandPrize2009\_BPC\_BellKor.pdf

Bipin, K. (2018). *Robot Operating System Cookbook.* Packet Publishing.

Das, S., Dai, R., Koperski, M., Minciullo, L., Garattoni, L., Bremond, F., & Francesca, G. (2019). Toyota Smarthome: Real-World Activities of Daily Living. *International Conference on Computer Vision* (pp. 833-842). Seoul, Korea: IEEE. doi:10.1109/ICCV.2019.00092

Delgado, R. (2021). A semi-hard voting combiner scheme to ensemble multi-class probabilistic classifiers. *Applied Intelligence, 1*, 1-25. doi:10.1007/s10489-021-02447-7

Ferati, M., Kurti, A., Vogel, B., & Raufi, B. (2016). Augmenting requirements gathering for people with special needs using IoT. *International Workshop on Cooperative and Human Aspects of Software Engineering* (pp. 48-52). ACM. doi:10.1145/2897586.2897617

García-Pérez, M. A. (2012). Statistical conclusion validity. *Frontiers in Psychology, 3*. doi:10.3389/fpsyg.2012.00325

Koreshoff, T., Robertson, T., & Leong, T. (2013). Internet of Things: a review of literature and products. *CHI’13, November 25 - 29 2013, Adelaide, Australia*.

Litomisky, K. (2012). *Consumer RGB-D Cameras and their Applications*. Retrieved from University of California: https://alumni.cs.ucr.edu/~klitomis/files/RGBD-intro.pdf

Mickens. (2018, August 16). *Why Do Keynote Speakers Keep Suggesting That Improving Security Is Possible?* Retrieved from YouTube: https://www.youtube.com/watch?v=ajGX7odA87k

Morris, J. (2008). *Disability research and policy: current perspectives.* Lawrence Erlbaum Associates.

Muhsin, A., Munyogwa, M., Kibusi, S., & Seif, S. A. (2020). Poor knowledge on elderly care despite positive attitude among nursing students in Zanzibar Island: findings from a cross-sectional study. *BMC Nursing, 19*(1), 1-8. doi:10.1186/s12912-020-00488-w

Parker, R. (1993). Threats to the validity of research. *Rehabilitation Counseling Bulletin, 36*(3), 130-138. Retrieved from https://search-ebscohost-com.proxy1.ncu.edu/login.aspx?direct=true&db=eric&AN=EJ458938&site=eds-live

Piirainen, K., & Gonzalez, R. (2013). Constructive Synergy in Design Science Research: A Comparative Analysis of Design Science Research and the Constructive Research Approach. *Liiketaloudellinen Aikakauskirja, 3*(4), 206-234. Retrieved from https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=bth&AN=95116694&site=eds-live

Silvestrini, R. P., & Sammito, G. (2012). Design of Experiments for Information Technology Systems. *Defense AT&L, 41*(5), 30-35. Retrieved from https://search-ebscohost-com.proxy1.ncu.edu/login.aspx?direct=true&db=bth&AN=80409129&site=eds-live

Singh, T., & Vishwakarma, D. (2018). *Human Activity Recognition in Video Benchmarks: A Survey.* Springer.

Snee, R. (2015). A practical approach to data mining. *Quality Engineering, 27*(4), 477-487. doi:10.1080/08982112.2015.1065322

Stanford Artificial Intelligence Laboratory et al. (2018). *Robotic Operating System*. Retrieved from https://www.ros.org

US Bureau of Labor Statistics. (2020, May). *Registered Nurses*. Retrieved from US Bureau of Labor Statistics: https://www.bls.gov/ooh/healthcare/registered-nurses.htm

Wachter, S. (2018, June). Normative challenges of identification in the Internet of Things: privacy, profiling, discrimination, and the GDPR. *Computer Law & Security Review, 34*(3), 436-449. doi:https://doi.org/10.1016/j.clsr.2018.02.002

Wilson, C. (2017). *Caring for people with dementia: a shared approach.* Los Angeles, California: SAGE.

Wood, T. (n.d.). *What is the F-score*? Retrieved from Deep AI: https://deepai.org/machine-learning-glossary-and-terms/f-score

Yang, M., Min, G., Yang, G., & Li, Z. (2014). Software rejuvenation in cluster computing systems with dependency between nodes. *Computing, 96*, 503–526. doi:10.1007/s00607-014-0385-x

# Annotated Bibliography

Examining the Northcentral University (NCU) Library with search terms such as elderly care, IoT, and video health monitoring uncovers several industry-wide trends.

## Human Activity Recognition

Amazon Go enables customers to purchase goods from physical stores without requiring cashiers (Amazon, 2021). Their solution uses Deep Neural Network (DNN) algorithms that process real-time video streams. Wankdhede et al. (2018) assessed the system’s sophistication through a series of shoplifting test cases. Their malicious attempts to steal items failed, proving that real-time video monitoring is an effective real-world tool. However, several critical changes are necessary before engineers can transpose the solution directly into a person’s home.

Action tracking systems require sophisticated machine learning algorithms that classify spatiotemporal movements. They typically follow a process that collects RGB+D (Color and Depth) camera frames and decodes them into 3-D space (Litomisky, 2012). Next, a series of filters must crop, resize, and align the principal subject. These steps are necessary because people can freely move around the room. Third, a feature extract process outputs matrices representing the body’s location, pose, and orientation. Finally, those tensors flow into Long-Term Short-Term (LTSM) algorithms that classify the movement into gestures.

Many general-purpose gesture detection libraries already exist for behaviors such as sporting events and other high-energy actions. However, daily indoor activity tends to be more subtle and nuanced (Das et al., 2019). This discrepancy limits reusability and creates the need for purpose-built training sets that sufficiently cover patient-specific actions. Furthermore, these custom datasets necessitate vast quantities of examples with accurate labels, or the machine learning model will produce unreliable results. As a general rule of thumb, each model parameter needs at least ten examples to avoid overfitting (Snee, 2015). Gesture models can quickly explode into hundreds or thousands of parameters which represents the 3D space plus time. Mechanisms exist for accelerating the process of building custom datasets (e.g., transfer learning). However, this is still an open research topic.

## Integrating IoT Systems

After the system decodes the subject’s actions, it needs to act upon that information. While some responses are relatively trivial (e.g., dispatch an ambulance), other reactions must utilize Cyber-Physical Systems (CPS) to cross the digital boundary. Building these integrations is challenging as it draws upon knowledge spanning networking, sensors, embedded systems, and related concepts. Furthermore, CPS devices lack standardization which impacts component reusability.

Elloumi et al. (2020) propose a Smarthouse Operating System (SOS) that provides core services such as identity management, system state, and message routing. Their blueprint also outlines several automation application profiles (e.g., heat management and fire detection). These abstract constructs enable developers to focus on their integrations value differentiation versus writing tedious generic code. Finally, the authors demonstrate the effectiveness of this approach using a CNC (Computer Numerical Control) machine to print a replica house (2 cubic feet).

## Enhancing Security

It can be helpful to think of identity within IoT as a profile of historical choices, stated preferences, user roles, and known associations (Wachter, 2018). When the device understands the user’s profile, the experience can be customized and produce more accurate predictions. However, the payment for access to these inferences and decision processes comes from personal information, such as calendars, contacts, and routines (Mickens, 2018). This trade creates privacy concerns that can be subtle and can go unnoticed for some time.

The monitoring system will collect intimate knowledge of its assigned patients and even capture private communications. Therefore, mechanisms must exist for protecting this information as it flows between different compute domains (e.g., local versus cloud providers). Yang et al. (2021) propose a Secure MultiParty Computation (SMC) model that locally encrypts sensitive images (e.g., faces). Afterward, they use Cheon-Kim-Kim-Song Homomorphic Encryption (CKKS HE) to predict the encrypted payload remotely (e.g., Alice versus Bob’s face). Ideally, the system minimizes the information that leaves the patient’s private network. However, these remote compute requirements are unavoidable in specific scenarios.

## Healthcare and Cloud

Software that uses cloud resources gains agility, elasticity, instantaneous provisioning, and cost management constructs. However, some businesses are reluctant to trust these environments entirely due to security concerns (Ali et al., 2015). Alternatively, systems engineers can deploy hardware appliances that bring cloud aspects into the home or medical facility. For example, Abdulameer et al. (2020) propose implementing a small replica house similar to Elloumi. Their solution uses various wearable device sensors, Arduino micro-controllers, and one Raspberry-PI. In addition, users can check their vitals and other health KPIs (Key Performance Indicators) through a web portal. Therefore, provisioning small on-premise appliances makes sense and handles scenarios such as failures at the ISP (Internet Service Provider). In addition, other benefits come from standardizing the control-plane versus assuming the patient’s personal computer is compatible and Always-On/Always Connected (AoAC).

Yi & Feng (2021) recently proposed a complete video-based injury rehabilitation solution that includes support for CPS and wearables. The authors leverage Carnegie Mellon University’s Open Pose library to map skeletal structures within images. Then, they publish this information with various sensor readings (e.g., smoke detectors) into a secure private cloud. Unlike Das et al. (2019), the authors use Dynamic Time Warping (DTW) to compare and categorize patients’ movements. Researchers use DTW to normalize time series and avoid discrepancies from action speeds (e.g., raising one’s hand within two versus four seconds).

Chen, Saiki & Nakamura (2020) state that monitoring low-intensity slow physical movements is challenging. These issues arise because training data is not broadly available due to researchers focusing on fast-paced sporting video by default. Their study uses PoseNet to track skeletal movements and predict activities. Additionally, they compare the resources necessary for Raspberry PI and a desktop computer to make those predictions. The researchers assess the model’s accuracy in terms of delta changes in the bounding box. This approach is unique and comes with several limitations. For instance, a person laying on the couch versus standing will have different dimensions. While this methodology is not directly usable, it provides another measurement strategy.

Chaing et al. (2011) propose a Uniform Markup Language (UML) model for collecting health care metadata from video sources. Their solution focuses on physiological information, such as the patient’s movements. The model also describes a storage structure for persisting the recordings. There are specific aspects from this study that are reusable. For instance, the authors propose service interfaces to several patient monitoring components (e.g., heart rate and oxygen levels). However, the video monitoring system is critically dependent on wearable technologies. This requirement makes the solution loosely relevant to the ESHOS project.

Nugroho, Harmanto & Al-Absi (2018) propose a deep learning model to assess a patients’ pain level. Their solution uses facial expressions from fourteen people that train both OpenFace and FaceNet topologies. The researchers claim that they can predict with 93% accuracy a person’s level of pain. This capability is helpful within home monitoring systems as a mechanism to assess medication levels. The ESHOS solution could introduce similar capabilities to improve patient care. For instance, the system could emit a pain frequency Key Performance Indicator (KPI) as part of the patient’s scorecard. The scorecard would then grant reassurance to family members that sufficient attention is available.