

Food for thought: Ethical considerations of user trust in computer vision

Kaylen J. Pfisterer^{*,1,3}, Jennifer Boger^{2,3}, and Alexander Wong^{1,3}

¹ Vision and Image Processing Research Group, University of Waterloo, Waterloo, ON, Canada

² Intelligent Technologies for Wellness and Independent Living Lab,
University of Waterloo, Waterloo, ON, Canada

³ Schlegel-UW Research Institute for Aging, Waterloo, ON, Canada

*kpfisterer@uwaterloo.ca

Abstract

In computer vision research, especially when novel applications of tools are developed, ethical implications around user perceptions of trust in the underlying technology should be considered and supported. Here, we describe an example of the incorporation of such considerations within the long-term care sector for tracking resident food and fluid intake. We highlight our recent user study conducted to develop a Goldilocks' quality horizontal prototype designed to support trust cues in which perceived trust in our horizontal prototype was higher than the existing system in place. We discuss the importance and need for user engagement as part of ongoing computer vision-driven technology development and describe several important factors related to trust that are relevant to developing decision-making tools.

1. Introduction

Ethical considerations around users' perceptions of trust in technology is a crucial but often neglected aspect when conducting computer vision research in areas that can have a major impact on societal well-being. This was highlighted in our recent work developing a computer vision application targeted at addressing malnutrition in older adults living in long-term care (LTC). In particular, our Automated Food Imaging and Nutrient Intake Tracking (AFINI-T) system is an end-to-end system driven by computer vision. For context, approximately 25% of the population worldwide has some form of malnutrition (*e.g.*, iron deficiency, zinc deficiency, food-insecurity malnutrition) [32, 16, 20] and this risk increases with age [5]. In LTC malnutrition prevalence in older adult residents is estimated to be 44% [15]. Food and fluid intake tracking are mandated in LTC for those at-risk for malnutrition [27], however existing methods rely on subjective, inaccurate, and typically retrospective meth-

ods [1, 7]. We aim to support improving quality of care, and thereby residents' quality of life, by harnessing computer vision to create an objective automation of this process in this historically subjective, paper-based application.

AFINI-T is a computer vision-based system comprised of four subsystems that perform cascading tasks: (1) *Optical*: images are acquired using an integrated RGB-D camera; (2) *Food segmentation*: discrimination between foods, the plate and the background; (3) *Food volume estimation*: feature extraction to describe each food segment and classification of each food segment to inform the (4) *Food intake and nutrient estimation*: draw inferences of the nutritional information for each food segment and summarize nutrient intake across a meal.

Others have made advances on aspects of this problem however, relatively few end-to-end systems have been developed [21, 26, 30] and these approaches have focused on an individual tracking their own intake for the purpose of weight loss or maintenance. While supporting individual nutritional tracking is certainly needed with tripled obesity rates since 1975 [34], this approach is incongruous with the needs in LTC. The application of a computer-vision based solution in LTC where caregivers and dietitians assess and report residents' food and fluid intake needs has yet to be addressed. Here, the environment is relatively constrained as menu items are relatively set, portion sizes are regulated, and many individuals are monitored by caregivers at each meal.

To address this need, we conducted a participatory iterative design sprint with representative end-users (*e.g.*, personal support workers, dietitians etc.) to inform the needs and problem space, refine the design, and evaluate the resulting early stage user-interface prototype for usability [29]. User centred and participatory action design methodologies can help overcome barriers to uptake [18, 12, 25] and early on user studies and ethical implication consideration of technology are major predictors for user acceptance [12]. While the general population may distrust

artificial intelligence and related fields including computer vision [28, 31], we present a different perspective of positive trust and the need to be responsible with that trust because our designs will be taken at face value. Here, we present one finding that stood out: higher perceived trust in our low-fidelity prototype than in the existing system. Using data from our prototype development, this paper discusses several ethical considerations within the context of applied computer vision techniques and systems in an interdisciplinary environment.

2. Methods

A 6-stage iterative participatory design sprint was developed and executed to address the importance and need for user engagement as part of ongoing computer vision driven technology. These six stages included: (1) Design Ideation, (2) Reflect and Storyboard, (3) Storyboard Critiques, (4) Design of the Goldilocks Quality Horizontal Prototype, (5) Usability Assessment, and (6) Final Validation. User feedback and output from each stage was incorporated to inform each proceeding stage. A total of 38 participants and advisors representing 15 distinct roles (*e.g.* personal support worker, nurse, and dietitian) were engaged in the design sprint. Subjective workload (Raw Task Load Index), subjective usability scales, and a modified Ravden checklist were used to assess project advisors perceptions of the AFINI-T system prototype compared with the existing method of food and fluid intake charting as described in [29]. Here we focus on the output 5 project advisors' evaluation of user perceptions of trust in the AFINI-T system compared to the existing method of food and fluid intake charting. A subset of Jian *et al.* [13]'s tool was used to capture perceptions related to deception, wariness, confidence, dependability, reliability, trust and familiarity with the system [11]. Statements were comprised of 7-point Likert scales ranging from not at all (1) to extremely (7). Responses were re-categorized from a 7-point Likert scale to "No", "Neutral", and "Yes" to summarize trends; the original 7-point Likert scale ratings were used for calculating a two-tailed t-test assuming unequal variances [9, 24] to compare the existing electronic paper-based system and the AFINI-T prototype [13].

3. Results

Consistent with our previously reported high usability (SUS score of 89.2), and significantly higher perceived performance with the AFINI-T prototype than the existing system ($P < .05$) in [29], Table 1 indicates there is low trust in the existing system (55% of respondents do not trust the system, $n=11$), the AFINI-T system was perceived to be more trustworthy, and generally, AFINI-T system trust ratings were opposite and more positive compared to the exist-

ing system for food and fluid intake charting (Table 1). For example, advisors rated the AFINI-T system as less deceptive (deceptive "yes: 17% AFINI-T, 45% existing system), less wary of the system (wariness "yes: 0% AFINI-T, 50% existing system), and more confident in the AFINI-T system (confident "yes: 83% AFINI-T, 18% existing system). As shown in Table 1 using a two-tailed t-test assuming unequal variances [9, 24] indicate that these three scores of deceptiveness, wariness, and confidence were significantly different between systems all with p -values < 0.05 . While not statistically significant, there were additionally higher ratings for dependability, reliability, and familiarity with the AFINI-T system prototype compared to the existing system. The statement regarding "I can trust" the system was higher for AFINI-T system prototype than to the existing system and this difference approached significance ($p=0.08$).

4. Discussion and Conclusion

We were particularly interested in trust as it pertains to the ability to introduce a new level of automation in this field by leveraging computer vision techniques. Several factors described by others that influence trust relevant to developing computer-vision-based applications are: the type of system, how complex the system is, how the system will be used, and the cost-benefit of using the system (*e.g.*, high risk or low risk) [11]. However, in the case of the AFINI-T system, we must be careful since this is a tool to support care in a vulnerable setting (older adults). As such, it must be clear that decision making rests with humans and reinforce that the AFINI-T system is a decision-making aid. Specifically, when automation comes with great benefit and fewer risks, people tend to increase reliance on automation; even under high-risk scenarios if the level of automation is low [11, 19] and humans tend to view automated tools as more accurate than humans [19]. This contributes to over-trust [17] and may be more important to consider than initially thought given readiness to accept the prototype-level technology and high trust ratings for even for the prototype that is not fully developed.

Our data collection strategy was grounded in theory, guided by several conceptual frameworks, and grounded expertise to complement the interdisciplinary and complexity of the problem space (*e.g.*, [4, 6]). We also borrowed from transdisciplinary research that "explicitly recognizes the value of partnerships and the different stakeholders along with their roles in facilitating and supporting innovation" [4]. In the AFINI-T design this was reflected through recruiting diverse and multidisciplinary project advisors. Our participatory iterative design sprint was employed to work directly with representative end-users to understand and incorporate their perspective and concerns to consider and appropriately support trust from multiple angles. By doing so, we aimed to support trust cues, cred-

Table 1. Advisors' perceived trust of existing food/fluid intake system (Existing) to the AFINI-T prototype (AFINI-T). 7-point Likert scale ratings were condensed to "No" for Likert ratings 1-3, "neutral" for a rating of 4, and "yes" for ratings 5-7.

Trust Statement	System	"No" n (%)	"Neutral" n (%)	"Yes" n (%)	Total n	t, df P-value
The food/fluid intake system is deceptive.	Existing	1 (9%)	4 (45%)	4 (45%)	11	3.45, df=11.2 (P = .005)
	AFINI-T	5 (83%)	0 (0%)	1 (17%)	6	
I am wary of the food/fluid intake system.	Existing	3 (30%)	2 (20%)	5 (50%)	10	3.34, df = 12.5 (P = .006)
	AFINI-T	5 (83%)	1 (17%)	0 (0%)	6	
I am confident in the food/fluid intake system.	Existing	5 (45%)	4 (36%)	2 (18%)	11	2.42, df = 8.03 (P = .042)
	AFINI-T	1 (17%)	0 (0%)	5 (83%)	6	
The food/fluid intake system is dependable.	Existing	4 (40%)	3 (30%)	3 (30%)	10	1.55, df = 6.80 (P = .166)
	AFINI-T	1 (17%)	0 (0%)	5 (83%)	6	
The food/fluid intake system is reliable.	Existing	5 (50%)	2 (20%)	3 (30%)	10	1.07, df = 7.83 (P = .319)
	AFINI-T	1 (17%)	1 (17%)	4 (67%)	6	
I can trust the food/fluid intake system.	Existing	5 (45%)	4 (36%)	2 (18%)	11	2.00, df = 8.29 (P = .080)
	AFINI-T	1 (17%)	0 (0%)	5 (83%)	6	
I am familiar with the food/fluid intake system.	Existing	1 (9%)	4 (36%)	5 (55%)	11	1.31, df=14.9 (P = .210)
	AFINI-T	0 (0%)	0 (0%)	6 (100%)	6	

ibility [8, 10], and to adhere to best practices for user interface design [2, 3, 14, 22, 23, 33]. This translated to incorporating ease of navigation (reinforced by click-saving features), use of good visual design elements by on-screen chunking of relevant information, aiming for an overall professional look, supporting search ability as well as smart guidance through transactions and overall, working directly with users to provide appropriate and useful content.

Particularly in computer vision research building intelligent systems with end-users engaged, we should be aware of how the perception of the technology changes through this process as suggested through this work. Specifically, we must be wary of potential over-trust in the system, which is when trust exceeds system capabilities [17]. Of interest in this application, the risk for over-trust is higher when building tools that will be used in high stress environments due to time constraints and have potentially large benefits over the existing method (*e.g.* time saving, improved quality of care) [11]. Regarding the receptivity to the AFINI-T system, advisors commented "Having data accessible opens door to how can be analyzed" and "This would save a lot of time especially if individualized. [The % daily value of nutrients is a] proportional calculation [that is currently] a manual process." In contrast, perhaps contributing to relatively low trust in the existing system, several users articulated that the quality of the existing method for data collection is not helpful at prioritizing resident referrals to dietitians. When considered in the LTC environment where being short-staffed is the norm, this may in part explain why trust ratings in the AFINI-T prototype were high and highlights the importance of these design considerations.

While fast-paced and time intensive, the results gathered in this study indicate it *is* possible to design AI/ML/CV ap-

plications with end-users and can result in high receptivity and trust in these technologies. Our experience corroborates that engaging with end-users throughout the process as collaborators enhanced a more comprehensive understanding of the problem space [4]. Moving forward, we must reassess trust in the fully functional system and probe more deeply into factors contributing to trustworthiness. More generally, as designers and researchers developing accountable computer vision systems we must be the first line of defense and should consider our moral obligation for ensuring accuracy, system reliability, and understanding the psychological effects of using trust cues to enhance usability. (Funding: Canada Research Chairs program; NSERC Postgraduate Scholarship-Doctoral)

References

- [1] Y. N. Andrews and V. H. Castellanos. Development of a method for estimation of food and fluid intakes by nursing assistants in long-term care facilities: a pilot study. *Journal of the American Dietetic Association*, 103(7):873–877, 2003.
- [2] Apple Inc. Human interface guidelines: Navigation bars. <https://developer.apple.com/design/human-interface-guidelines/ios/bars/navigation-bars/>, (accessed: 04.08.2018).
- [3] N. Babich. Button UX design: Best practices, types and state. <https://uxplanet.org/button-ux-design-best-practices-types-and-states-647cf4ae0fc6>, (accessed: 04.08.2018).
- [4] J. Boger, P. Jackson, M. Mulvenna, J. Sixsmith, A. Sixsmith, A. Mihailidis, P. Kontos, J. Miller Polgar, A. Grigorovich, and S. Martin. Principles for fostering the transdisciplinary development of assistive technologies. *Disability and Rehabilitation: Assistive Technology*, 12(5):480–490, 2017.

- [5] S. Brownie. Why are elderly individuals at risk of nutritional deficiency? *International Journal of Nursing Practice*, 12(2):110–118, 2006.
- [6] E. C. Carr, J. N. Babione, and D. Marshall. Translating research into practice through user-centered design: An application for osteoarthritis healthcare planning. *International Journal of Medical Informatics*, 104:31–37, 2017.
- [7] V. H. Castellanos and Y. N. Andrews. Inherent flaws in a method of estimating meal intake commonly used in long-term-care facilities. *Journal of the American Dietetic Association*, 102(6):826–830, 2002.
- [8] C. L. Corritore, B. Kracher, and S. Wiedenbeck. On-line trust: concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, 58(6):737–758, 2003.
- [9] A. Field. *Discovering statistics using IBM SPSS statistics*. sage, 2013.
- [10] B. Fogg and H. Tseng. The elements of computer credibility. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 80–87. ACM, 1999.
- [11] K. A. Hoff and M. Bashir. Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3):407–434, 2015.
- [12] M. Ienca, T. Wangmo, F. Jotterand, R. W. Kressig, and B. Elger. Ethical design of intelligent assistive technologies for dementia: a descriptive review. *Science and Engineering Ethics*, 24(4):1035–1055, 2018.
- [13] J.-Y. Jian, A. M. Bisantz, and C. G. Drury. Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, 4(1):53–71, 2000.
- [14] Justinmind. 7 rules for mobile UI button design. <https://uxplanet.org/7-rules-for-mobile-ui-button-design-e9cf2ea54556>, (accessed: 04.08.2018).
- [15] H. H. Keller, N. Carrier, S. E. Slaughter, C. Lengyel, C. M. Steele, L. Duizer, J. Morrison, K. S. Brown, H. Chaudhury, M. N. Yoon, et al. Prevalence and determinants of poor food intake of residents living in long-term care. *Journal of the American Medical Directors Association*, 18(11):941–947, 2017.
- [16] D. B. Kumssa, E. J. Joy, E. L. Ander, M. J. Watts, S. D. Young, S. Walker, and M. R. Broadley. Dietary calcium and zinc deficiency risks are decreasing but remain prevalent. *Scientific Reports*, 5:10974, 2015.
- [17] J. D. Lee and K. A. See. Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1):50–80, 2004.
- [18] M. Leo, G. Medioni, M. Trivedi, T. Kanade, and G. M. Farinella. Computer vision for assistive technologies. *Computer Vision and Image Understanding*, 154:1–15, 2017.
- [19] J. B. Lyons and C. K. Stokes. Human–human reliance in the context of automation. *Human Factors*, 54(1):112–121, 2012.
- [20] S. McGuire. State of food insecurity in the world 2015: Meeting the 2015 international hunger targets: Taking stock of uneven progress. Rome: FAO, 2015. <http://www.fao.org/3/a-i4646e.pdf>.
- [21] A. Meyers, N. Johnston, V. Rathod, A. Korattikara, A. Gorban, N. Silberman, S. Guadarrama, G. Papandreou, J. Huang, and K. P. Murphy. Im2calories: towards an automated mobile vision food diary. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1233–1241, 2015.
- [22] J. Nielsen. 10 usability heuristics for user interface design. *Nielsen Norman Group*, 1(1), 1995.
- [23] D. Norman. *The design of everyday things: Revised and expanded edition*. Constellation, 2013.
- [24] G. R. Norman and D. L. Streiner. *Biostatistics: The bare essentials*. PMPH-USA, 2008.
- [25] P. Novitzky, A. F. Smeaton, C. Chen, K. Irving, T. Jacquemard, F. O'Brien, D. OMáthúna, and B. Gordijn. A review of contemporary work on the ethics of ambient assisted living technologies for people with dementia. *Science and Engineering Ethics*, 21(3):707–765, 2015.
- [26] K. Okamoto and K. Yanai. An automatic calorie estimation system of food images on a smartphone. In *Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management*, pages 63–70. ACM, 2016.
- [27] Ontario Association of Non-Profit Homes and Services for Seniors. OANHSS Nutrition and Hydration Program - Policy, Procedures and Training Package. <http://ltctoolkit.rnao.ca/resources/policies-and-procedures/oanhss-nutrition-and-hydration-program-policy-procedures-and-train>, (accessed: 17.02.2019).
- [28] Pega. What consumers really think about AI: A global study. Technical report, Pegasystems, 2017. <https://www.ciosummits.com/what-consumers-really-think-about-ai.pdf>.
- [29] K. Pfisterer, J. Boger, and A. Wong. Prototyping the automated food imaging and nutrient intake tracking (AFINT) system: A modified participatory iterative design sprint. *JMIR Hum Factors*, 6(2):e13017, May 2019.
- [30] P. Pouladzadeh, S. Shirmohammadi, and A. Yassine. You are what you eat: So measure what you eat! *IEEE Instrumentation & Measurement Magazine*, 19(1):9–15, 2016.
- [31] pwc. Accelerating innovation: How to build trust and confidence in AI. Technical report, pwc, 2017. <https://www.pwc.co.uk/audit-assurance/assets/pdf/responsible-artificial-intelligence.pdf>.
- [32] A. Rojer, H. Kruizenga, M. C. Trappenburg, E. Reijnerse, S. Sipilä, M. V. Narici, J. Hogrel, G. Butler-Browne, J. S. McPhee, M. Pääsuke, et al. The prevalence of malnutrition according to the new espn definition in four diverse populations. *Clinical Nutrition*, 35(3):758–762, 2016.
- [33] B. Schneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmquist, and N. Diakopoulos. *Designing the user interface: strategies for effective human-computer interaction*. Pearson, 2016.
- [34] World Health Organization. Obesity and overweight. <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>, (accessed: 02.21.2019).