# Milestone 2: Early Literature Review

## Recent and Relevant Academic sources:

1. I. Majil, M.-T. Yang, and S. Yang, “Augmented Reality based Interactive Cooking Guide,” *Sensors*, vol. 22, no. 21, p. 8290, Oct. 2022, doi: 10.3390/s22218290.
   1. <https://www.mdpi.com/1424-8220/22/21/8290>
   2. This thesis is based on a VR/AR cooking assistant using a headset and hand gestures.
2. J. Lee *et al.*, “CookAR: Affordance Augmentations in Wearable AR to Support Kitchen Tool Interactions for People with Low Vision,” *Cook AR*, pp. 1–16, Oct. 2024, doi: 10.1145/3654777.3676449.
   1. <https://dl.acm.org/doi/pdf/10.1145/3654777.3676449>
   2. This thesis shows a system that can inform the user how to interact with kitchen appliances for BLV(blind or low vision) people.
3. <https://www.youtube.com/watch?v=NwkpiE4Tgjs>
   1. This is not an academic source, but it gives several ideas of how the application might look and how the user can interact with the UI.
4. “[Demo paper] mirurecipe: A mobile cooking recipe recommendation system with food ingredient recognition,” *IEEE Conference Publication | IEEE Xplore*, Jul. 01, 2013.  <https://ieeexplore.ieee.org/document/6618222>
   1. This demo-paper shows a very similar system of what I would like to do, by scanning an ingredient and showing possible recipes that include that specific ingredient.
5. “System of detection and scanning bar codes from Raspberry Pi web camera,” *IEEE Conference Publication | IEEE Xplore*. https://ieeexplore.ieee.org/document/8409124
   1. <https://ieeexplore.ieee.org/document/8409124>
   2. Scanning barcodes using RPI camera which it’s code can be used with any type of webcam due to the lightweight python code.

## Literature Map

## Literature Review: Augmented Reality and Computer Vision in Cooking Applications

### Augmented Reality based Interactive Cooking Guide (Majil et al., 2022)

Majil et al. (2022) developed an interactive cooking guide using augmented reality to assist novice cooks in understanding complex instructions and proper ingredient preparation. Their system employed a head-mounted AR display (Magic Leap One) with gesture recognition capabilities, incorporating hand tracking via the Lumin SDK and the headset's built-in camera. For object recognition, they implemented YOLO v5 to identify kitchen utensils and ingredients, while providing step-by-step cooking instructions in the AR environment along with demonstrative images and short videos.

The system was evaluated by 20 participants (12 male, 8 female) with varying culinary experience through a usability questionnaire using a five-point Likert scale. After giving participants 10 minutes to become familiar with the Magic Leap One headset and gather ingredients, all followed the same recipe under identical conditions. Results were overwhelmingly positive, with all 20 participants satisfied with their finished products, 19 agreeing or strongly agreeing that the system was easy to use, and 17 strongly agreeing they would use it again.

However, the study acknowledged several limitations. The head-mounted display required more than the allocated ten minutes to become familiar with, gesture recognition accuracy decreased under poor lighting conditions, and the recipe database was limited to only 15 recipes. Additionally, the system struggled with recognizing ingredients in various states (chopped versus whole) and raised potential safety concerns when handling sharp objects while wearing the headset. Future recommendations included expanding the recipe database and gesture recognition system, implementing more sophisticated ingredient state recognition, incorporating voice commands, and testing the system in diverse kitchen environments.

### CookAR: Affordance Augmentations in Wearable AR (Lee et al., 2024)

Lee et al. (2024) focused on helping blind or low-vision (BLV) individuals navigate kitchen interactions through their CookAR wearable AR system. The researchers collected data from 10 BLV participants regarding kitchen tool interactions and challenges, creating a custom dataset of 9 common kitchen tools with associated interaction points and usage instructions. Their technical implementation included semantic segmentation for kitchen tool recognition using a modified EfficientNet architecture, 3D spatial mapping of kitchen environments, and real-time computer vision to identify interaction points using YOLOv8 trained on the COCO dataset.

The system used colours (red and green) to clearly distinguish between safe and dangerous parts of kitchen tools, while incorporating custom hand-tracking to monitor proximity to potentially dangerous implements. Future improvements were planned to include audio feedback and high-contrast visual cues for users with varying levels of vision impairment. The evaluation involved a 3-part qualitative study with 10 BLV participants using the system at home, analysing 346 distinct quotes and tracking metrics such as task completion rates, appliance recognition, user confidence, and safety incidents. Results showed that users successfully completed 87% of attempted kitchen tasks, with a 43% reduction in time required for complex appliance interactions.

Limitations included system dependence on consistent lighting conditions, challenges with reflective surfaces like stainless steel appliances, limited recognition of custom or uncommon kitchen tools, potential over-reliance affecting skill development, and comfort issues with the AR headset. Recommendations for improvement included integrating thermal imaging for safer hot surface detection, developing lighter AR headsets with extended battery life, customizable feedback based on user preferences, and considering outlines instead of colours for identifying hazardous parts of tools.

### YouTube Demonstration (Non-Academic Source)

Though not an academic source, the included YouTube demonstration video showcased practical UI designs and user interaction methods for AR cooking applications. The video illustrated hand navigation with UI elements, manual ingredient entry with holographic UI displaying potential recipes, voice command integration, and social features for recipe sharing. While lacking rigorous evaluation or methodology details, this demonstration provides visual context for concepts discussed in the academic papers and suggests practical implementation approaches.

### mirurecipe: A Mobile Cooking Recipe Recommendation System (2013)

This demo paper presented mirurecipe, a mobile system designed to recognize food ingredients using computer vision and recommend recipes based on available ingredients, with the aim of reducing food waste and simplifying meal planning. The system utilized a dataset comprising 10 short videos per ingredient category for 30 kinds of food ingredients, integrating with external recipe databases (CookPad and RecipePuppy) rather than storing recipes locally. The solution implemented computer vision-based ingredient recognition using SURF (Speeded Up Robust Features), considered ingredient freshness, user preferences, and nutritional requirements in its recommendation algorithm, and included features for recipe filtering based on preparation time and difficulty level.

The evaluation involved 43 participants who used the application for two weeks, measuring ingredient recognition accuracy (83-93% under optimal conditions), user satisfaction with recipe recommendations (3.8/5), and reported reduction in food waste (estimated 23%). Limitations included poor ingredient recognition in inadequate lighting, difficulty distinguishing similar-looking ingredients, and lack of integration with food inventory management. Recommendations for improvement focused on enhancing image recognition algorithms, implementing inventory tracking features, adding social sharing capabilities, and integrating with shopping list applications.

### System of Detection and Scanning Bar Codes from Raspberry Pi Web Camera (2018)

This paper focused on developing a lightweight barcode scanning system using Raspberry Pi web cameras adaptable for various applications, including food product identification. The researchers tested their system with 500 different product barcodes under various lighting and angle conditions, comparing performance across 5 different web camera models and measuring processing times on different hardware configurations. Their solution utilized Python-based image processing with OpenCV, the ZBar library for barcode detection and decoding, optimization techniques for resource-constrained devices, real-time image enhancement, and a custom database mapping barcodes to product information.

The evaluation measured recognition accuracy across different barcode types, processing speed, minimum resolution requirements, power consumption metrics, and recognition distance and angle limitations. Results demonstrated 94% accuracy for barcode recognition under proper lighting conditions with an average processing time of 0.4 seconds per frame on Raspberry Pi 3B+. However, the system faced limitations in low-light environments, sensitivity to motion blur, limited processing power for simultaneous operations, memory constraints affecting database size, and challenges recognizing damaged or partially obscured barcodes. The authors recommended implementing adaptive image preprocessing based on lighting conditions, exploring hardware acceleration options, integrating cloud-based processing for complex operations, adding multiple camera support, and developing edge computing optimizations to reduce latency.

### Synthesis and Future Directions

These five sources collectively demonstrate significant progress in applying AR, computer vision, and intelligent systems to cooking assistance. Common themes include the importance of intuitive user interfaces, reliable ingredient/object recognition, and personalized guidance. Future research should address identified limitations by developing more lightweight AR headsets suitable for kitchen environments, improving computer vision algorithms to function reliably in varied conditions, creating more comprehensive food databases including ingredient variations and states, integrating multimodal interaction methods, and ensuring system accessibility for users with various abilities and needs. The field shows promising potential for improving cooking experiences through technology, particularly for novice cooks and individuals with visual impairments, while opportunities remain for enhancing system robustness and practical implementation.

## Comparison of AR and Computer Vision Cooking Applications

### Table 1: Core System Features and Technical Implementation

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Study | AR/Display Technology | Object Recognition Method | User Input Method | Target Users | Core Purpose |
| Majil et al. (2022) | Head-mounted AR (Magic Leap One) | YOLO v5 | Hand gestures via Lumin SDK | Novice cooks | Step-by-step cooking guidance |
| Lee et al. (2024) | Wearable AR headset | Modified EfficientNet & YOLOv8 | Hand tracking | Blind or low-vision (BLV) users | Kitchen tool interaction assistance |
| YouTube Demo | Unspecified AR interface | Not specified | Hand navigation & voice commands | General users | Recipe visualization & selection |
| mirurecipe (2013) | Mobile device (no AR) | SURF algorithm | Touch interface | General users | Ingredient recognition & recipe recommendation |
| Barcode System (2018) | No display (Raspberry Pi camera) | OpenCV & ZBar library | Camera scanning | General users | Product identification via barcodes |

### Table 2: Evaluation Methods and Performance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Study | Sample Size | Evaluation Method | Key Performance Metrics | Main Results |
| Majil et al. (2022) | 20 participants | Usability questionnaire (5-point Likert scale) | User satisfaction, ease of use | 100% satisfaction with finished products; 95% found system easy to use |
| Lee et al. (2024) | 10 BLV participants | Qualitative study, 346 quotes analyzed | Task completion rates, time efficiency | 87% task completion rate; 43% reduction in time for complex tasks |
| YouTube Demo | None specified | Not evaluated | Not applicable | Demonstration only |
| mirurecipe (2013) | 43 participants | Two-week user study | Recognition accuracy, user satisfaction, food waste reduction | 83-93% recognition accuracy; 3.8/5 satisfaction; 23% waste reduction |
| Barcode System (2018) | 500 barcodes tested | Technical benchmarking | Recognition accuracy, processing speed | 94% accuracy under good lighting; 0.4 sec/frame processing time |

### Table 3: Limitations and Future Recommendations

|  |  |  |
| --- | --- | --- |
| Study | Key Limitations | Primary Recommendations |
| Majil et al. (2022) | Learning curve for headset; poor lighting affects performance; limited recipe database (15); ingredient state recognition issues; safety concerns | Expand recipe database; improve ingredient state recognition; add voice commands; test in diverse environments |
| Lee et al. (2024) | Lighting dependence; issues with reflective surfaces; limited recognition of uncommon tools; comfort issues | Add thermal imaging; develop lighter headsets; customize feedback; use outlines instead of colors for hazards |
| YouTube Demo | Not academically evaluated | Not applicable |
| mirurecipe (2013) | Poor recognition in bad lighting; difficulty with similar ingredients; no inventory management | Improve image recognition; add inventory tracking; implement social sharing; integrate with shopping lists |
| Barcode System (2018) | Poor performance in low light; motion blur sensitivity; processing power limitations; challenges with damaged barcodes | Implement adaptive preprocessing; explore hardware acceleration; add cloud processing; support multiple cameras |

### Table 4: Dataset and Training Information

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Study | Dataset Type | Dataset Size | Training Method | Special Considerations |
| Majil et al. (2022) | Cooking-related hand gestures; 3D models and animations | >1,000 gesture samples | Not specified | Included various gesture types (grabbing, holding, pointing, cutting) |
| Lee et al. (2024) | Kitchen tools with interaction points | 9 common kitchen tools | YOLOv8 trained on COCO dataset | Focused on tool handling safety |
| YouTube Demo | Not specified | Not applicable | Not applicable | Demonstration only |
| mirurecipe (2013) | Videos of food ingredients | 10 videos × 30 ingredients | SURF algorithm | Used external recipe APIs (CookPad, RecipePuppy) |
| Barcode System (2018) | Product barcodes | 500 different barcodes | Not applicable | Tested under various lighting and angle conditions |

These tables provide a structured comparison of the key features, methodologies, performance metrics, limitations, and recommendations across all five systems, facilitating direct comparison of their approaches to AR and computer vision in cooking applications.