# Milestone 3: Research Methodology

### Determine whether the research aim, hypothesis and questions previously proposed need to be revised.

The research aim remains appropriate, maintaining its focus on exploring the effectiveness of Augmented Reality (AR) technology in enhancing cooking skills and boosting confidence in the kitchen, particularly for beginners. However, the hypothesis has been revised to reflect the current technical constraints. Specifically, the project will not include real-time guidance features—such as visually demonstrating how to cut an onion through a user interface in the AR display—due to limitations in AR functionality without the use of a dedicated VR headset.

### Propose a pipeline/plan for your research and prepare an illustration. Some sort of experimentation is needed for evaluation purposes. Use your early literature review as inspiration for candidate pipelines in your chosen subject area.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Phase 01: Initial Research and Setup | Phase 02: Data Acquisition | Phase 03: Experimentation | Phase 04: Evaluation | Phase 05: Improvements |
| 1.1 – Topic selection  1.2 – Literature Review  1.3 – System Setup | 2.3 – Requirement Mapping  2.1 – User Research  2.2 – Task Analysis | 3.3 – Focus Definition  3.2 – Feature Review  3.1 – Prototype design | 4.3 – Identify Gaps  4.2 – User Feedback  4.1 – Usability Testing | 5.1 – Improvement Research  5.2 – Prototype Update  5.3 – Final Review |

**Phase 01: Initial Research and Setup**

*Lays the groundwork for the project by identifying the topic, reviewing existing literature, and preparing development tools.*

* **1.1 Topic Selection** – Select the core research topic focused on AR’s role in enhancing beginner cooking skills.
* **1.2 Literature Review** – Review existing studies and technologies that apply AR in education and culinary training.
* **1.3 System Setup** – Set up development tools and frameworks required for building the AR prototype

**Phase 02: Data Acquisition**

*Gathers user needs and analyses cooking activities to define relevant features for the AR prototype.*

* **2.1 User Research** – Gather insights from students who do not have substantial experience in cooking activities to identify common challenges and learning needs.
* **2.2 Task Analysis** – Analyse basic cooking tasks to determine which are best supported by AR instructions.
* **2.3 Requirement Mapping** – Translate findings into technical and user experience requirements for the AR prototype.

**Phase 03: Experimentation**

*Involves developing, reviewing, and refining the initial prototype based on feasibility and research directions.*

* **3.1 Prototype Design** – build an initial AR interface to support basic cooking guidance trough overlays and prompts.
* **3.2 Feature Review** – Compare with similar AR tools and identify realistic features to implement withing limitations.
* **3.3 Focus Definition** – Define the research scope, excluding complex real-time features not feasible without VR gear.

**Phase 04: Evaluation**

*Tests the prototype with users and gathers feedback to measure effectiveness and identify areas for improvement.*

* **4.1 Usability Testing** – Observe users interacting with the prototype to assess ease of use and learning impact.
* **4.2 User Feedback** – Collect feedback to understand strengths, weaknesses, and areas for enhancement.
* **4.3 Identify Gaps** – Document limitations in design, functionality, and technology use that affect learning effectiveness.

**Phase 05: Improvements**

*Focus on refining the prototype and preparing it for final review based on user feedback and research findings.*

* **5.1 Improvement Research** – Investigate feasible ways to address identified limitations using available tools.
* **5.2 Prototype Update** – Integrate feedback-based changes to improve the AR cooking assistant experience.
* **5.3 Final Review** – Conduct a final evaluation and summarize limitations, strengths, and future enhancement opportunities.

### Proposed Research Methods

Data will be collected in a hybrid manner with qualitative data being gathered from user reviews and feedback whilst quantitative data will be collected via percentage values of precision of the device.

The qualitative data will be collected with a survey and face to face user feedback. This will all be recorded and organised in a useful manner to use this qualitative data and create an evaluation and note improvement possibilities.

Quantitative data will be gathered from repeating the ingredient detection accuracy from several tests and user results. For this to be a successful and reliable device, it has to have precision of 70% or higher. A number of images will be taken of a variety of ingredients to create the detection model for the YOLO ultralytics software to be as precise as possible. To calculate the precision percentage the following methods will be used:

1. IoU ( Intersection over Union) – Measures how much the predicted bounding box overlaps with the ground truth. **[1]**
   1. A prediction is considered correct if IoU >= 0.5 (common threshold)
2. Precision – Of all predicted boxes, how many were correct. **[2]**
3. Recall – Of all actual objects, how many were detected. **[3]**
4. mAp (means Average Precision) – The most common metric for object detection models. **[4]**
   1. It calculates the average precision (AP) for each class and then takes the mean.
   2. [mAp@0.5](mailto:mAp@0.5): uses IoU threshold of 0.5
   3. [mAP@[.5:.9.5](mailto:mAP@[.5:.9.5)]: averaged over multiple IoU thresholds (0.5 to 0.95 in 0.05 steps)

This can all be calculated with a simple script that uses a built in YOLO function that will give all the values that correspond to the results of the project.

Explanation of Terminology

Markings are made in the task 3 as follows: [n]

[1] – IoU Terminologies

* IoU – Intersection of union
  + Measures how much a predicted bounding box overlaps with the ground truth
  + If IoU is >= 0.5 – prediction counts as a True Positive (correct reading)
  + The higher the IoU, the better the bounding box match.
* Bounding Boxes
  + Bounding boxes are rectangles that define the position of an object in an image.
* Ground Truth Bounding Box
  + The ground truth is the actual label provided by a person that marks where the object really is in the image.
  + Ex: A model to detect cats is being used and in the image the human labelled the cats position or name with the box.
* Predicted Bounding Box
  + The predicted bounding box is what the model things the object location is
  + The goal is to make this predicted box match the ground truth box as close as possible
* The Overlap
  + This is where the IoU (intersection over union) comes in:
    - It checks how much the predicted box overlaps the ground truth box
    - Higher IoU = more accurate detection

[2] – Precision Terminologies

* True Positive (TP) – A correct detection
  + The model correctly detects an object, and the predicted bounding box overlaps well with the ground truth (usually IoU >= 0.5), and the predicted class is correct.
  + Example: A dog is labelled in the image, and the model predicted a “dog” with a bounding box that overlaps it well.
* False Positive (FP) – A wrong detection
  + This happens when the model detects something but its not correct.
  + This could be duplicate detection(multiple predictions on the same object), wrong class of objects, a bounding box that doesn’t match any real object.
  + Example: There is no cat in the image, but the model says “cat” with a box somewhere.
* False Negative (FN) - A missed detection
  + The model fails to detect an object that actually exists
  + Example: There’s a “car” in the image, but the model doesn’t detect it.
* True Negative (TN)
  + Not typically used in object detection
    - It refers to correctly identifying the absence of an object. Since object detection models are only looking for things, TNs are not calculated or relevant

How to get one single value from all the data representing the precision of the system:

A system called mAP(mean Average Precision) is used.

This is an already unified metric that combines precision and recall across different tresholds and classes.

It gives you a single score from 0 to 1 (or 0% to 100%) that reflects how well your model detects and classifies objects.

mAP computes a Precision-Recall curve for each class.

Then it calculates the area under this curve, which is called AP (Average Precision) for that class. Finally, it takes the mean of all Aps 🡪 mAP.

The formula is a s follows where *N* is the number of classes.

A mathematical equation on a black background

AI-generated content may be incorrect.