SmartPantry

Turning a Single Fridge Photo into Real Cooking Recommendations

Group Members: Stacy Che, Kexin Lyu, Samantha Wang, Zexi Wu (Allen), Tongrui Zhang
Date: December 8th, 2025
Class: DSAN 6700

Introduction

SmartPantry was designed around a simple, familiar problem: people often want to cook but lack the time or energy to figure out what they can make with what they already have. Our goal was to build an end to end system that accepts a single photo of a refrigerator and returns a ranked list of recipe ideas that match the detected ingredients. To achieve this, we built an image recognition module, an ingredient-level information retrieval engine, and a fully deployed web interface running on Hugging Face.

This project demonstrates the full lifecycle of a modern applied machine learning system. From data preprocessing and model benchmarking, to IR pipeline design, to containerized deployment.

Image Recognition

We used the Roboflow Smart Fridge dataset, which contains 1,474 well-annotated images of realistic refrigerator layouts. The dataset's variability in lighting, clutter, and angles made it suitable for a real world scenario rather than a controlled academic task.

We began with a ResNet50 multi-label classifier as a baseline. The task was straightforward: identify which ingredients appear in the image. The model, pretrained on ImageNet, provided a stable foundation for feature extraction but lacked spatial understanding. It could classify ingredients but could not localize them, and it performed inconsistently in crowded fridge scenes.

To address this, we transitioned to YOLOv8, a modern real time object detector. YOLOv8 provided stronger accuracy, better handling of overlapping objects, and immediate bounding box outputs that fed directly into our IR engine. The addition of single pass detection allowed fast inference suitable for deployment, and the anchor free architecture generalizes better on unusual refrigerator contents. This model became the backbone of the system.

Preprocessing and Normalization:

Ingredient text from recipe databases is messy. Our system removes measurement units, handles fractions, strips parentheses, and reduces ingredients to their interpretable root forms. We also implemented domain-specific stopword removal ("fresh," "organic") and ontological mapping of multi-word ingredients (e.g., "extra virgin olive oil" → "olive oil"). Lemmatization and stemming further standardized the final vocabulary.

Matching and Ranking:

The retrieval system first filters recipes that contain at least one matching ingredient. It then uses a combination of exact match, fuzzy match (≥75%), and ontology match to assemble a candidate set.

Missing ingredients are penalized, and matches are scored based on a continuous weight rather than binary logic. The top five recipes are then ranked according to ingredient coverage and similarity. This structure ensured that even imperfect YOLO predictions could still yield valid recipe suggestions, and that noisy ingredient text would not block relevant matches.

Deployment

We deployed the full system on Hugging Face Spaces, using a Docker container that bundled both the React frontend and FastAPI backend. The backend exposes a REST endpoint that receives a user-uploaded image, runs YOLOv8, preprocesses ingredients, retrieves matches, and returns a structured JSON response.

The frontend interacts only through API calls and displays recipe results cleanly, making the service approachable and lightweight. This architecture allowed us to deliver a fully functional live demo, and it creates a strong foundation for future extension.

Website & App Functionality

The SmartPantry web app delivers the full pipeline through a simple, intuitive interface. Users begin by uploading a photo of their fridge, which is immediately processed by the YOLOv8 detection model. The detected items populate an editable ingredient list, allowing users to add missing items or remove incorrect ones. Once the list is confirmed, the user clicks a single button to run the information-retrieval engine. The system then outputs a curated set of recipes, each with full ingredient lists, preparation steps, and estimated coverage of the user's available ingredients. This design keeps the interaction lightweight while still exposing the full capabilities of the underlying models.

Limitations

Although the system works reliably, several constraints remain. Image recognition still struggles with heavy occlusion or visually ambiguous items. Only 30 ingredient categories are supported, limiting the range of potential recipes. Ingredient normalization handles most cases but can be challenged by complex phrases. And while recipe matching works well at a structural level, it does not incorporate user preferences, cooking difficulty, or dietary constraints.

Conclusion

SmartPantry demonstrates a complete, practical pipeline for turning an unstructured fridge image into structured, actionable cooking advice. The combination of YOLOv8 for detection, our custom IR and normalization stages, and a fully deployed web application highlights how computer vision, information retrieval, and modern deployment tools can be brought together into a cohesive product. The project shows both the promise and the challenges of real-world ML applications, and it creates a pathway for expanding ingredient coverage, improving ontology depth, and incorporating personalization in future iterations.