Built-in

A GenAI-Based System for Verifying Compliance Between Floorplans Versions

Afik Aharon, Adi Haim, Liran Aichnboim

Project Review & Definition

MOTIVATION:

Comparing different versions of the same floorplan is difficult for buyers and may lead to financial and legal risks.

TASK:

detect meaningful content changes between two versions of the same floorplan image (before / after).

MODELS:

- Object Detection for structural understanding
- Segmentation for precise change localization
- Logistic Regression for classification

DATA:

- Synthetic data for before / after floorplan pairs
- Known change types and pixel-level masks

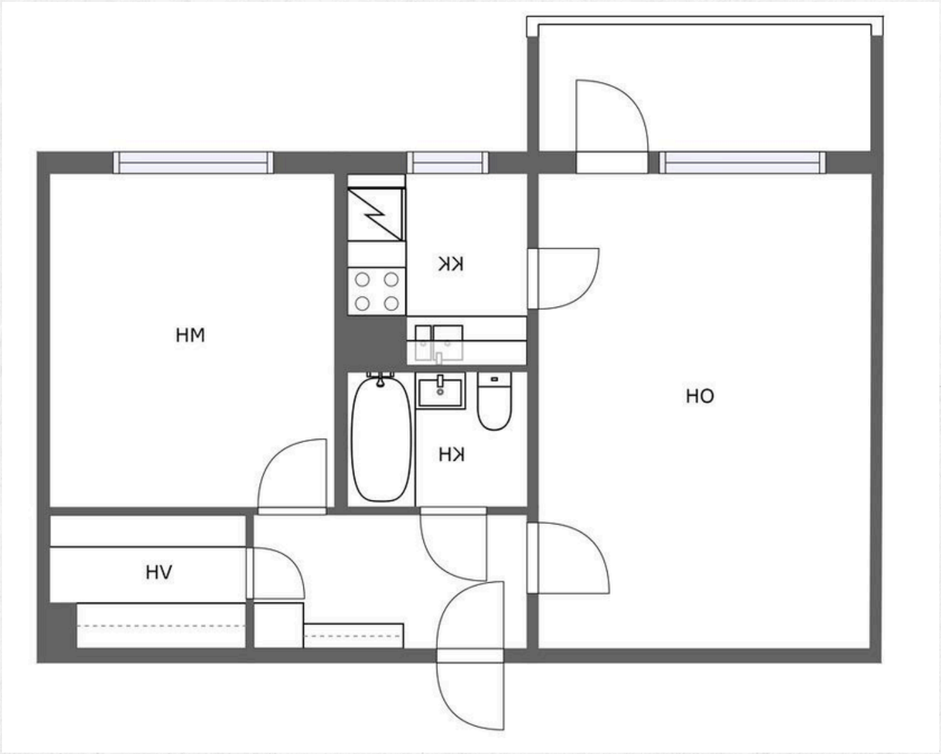
METRICS:

- Accuracy, Precision, Recall, F1-score
- Per change-type evaluation

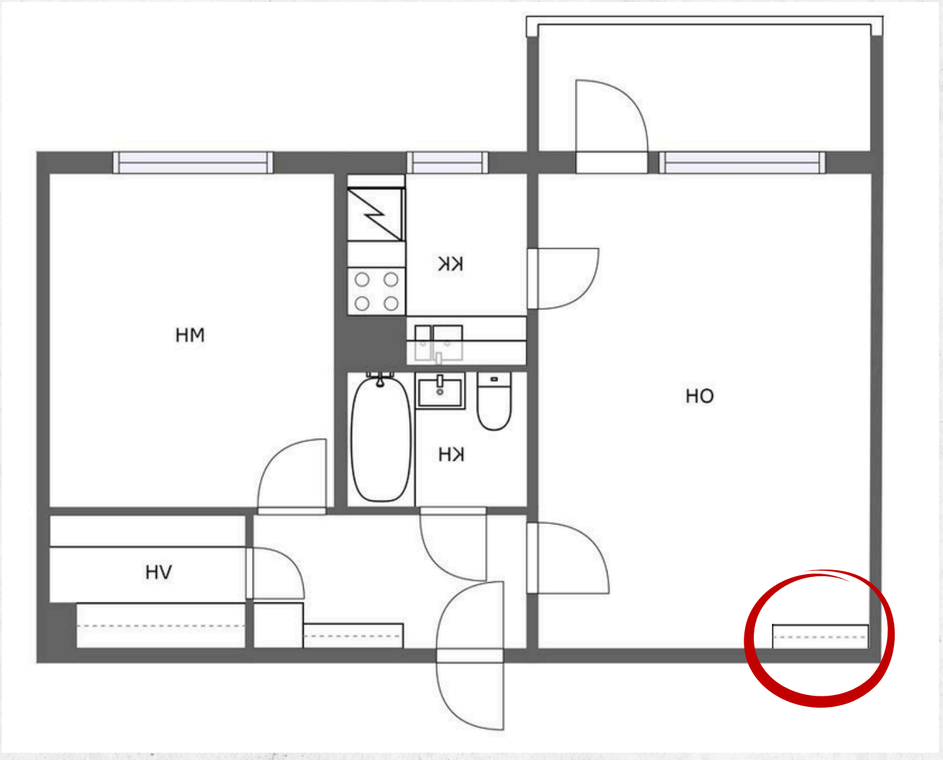
We generate the data needed to teach AI how to understand meaningful design changes.

Example 1

BEFORE:



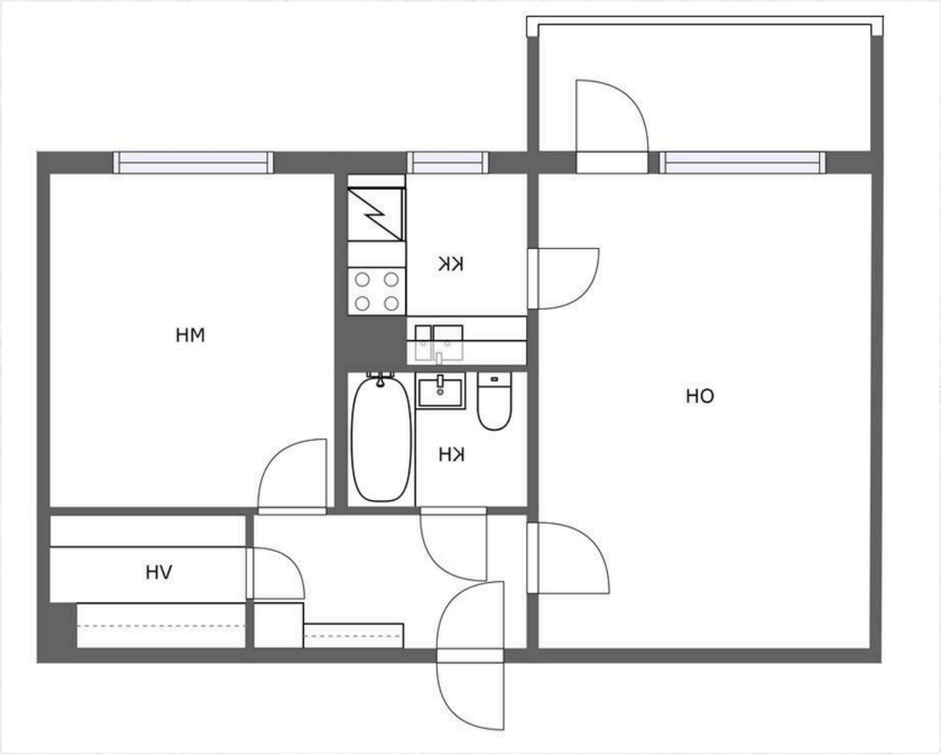
AFTER:



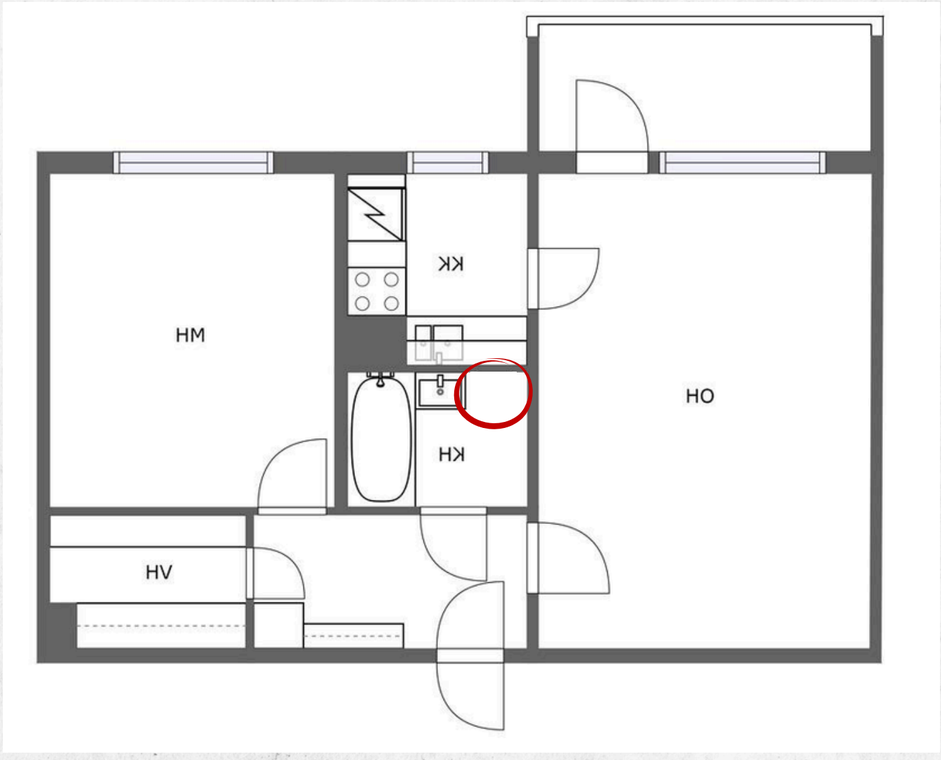
output: 0 – No substantial modification

Example 2

BEFORE:



AFTER:



output: 1 – substantial modification

Project achievements and novelty

PROJECT ACHIEVEMENTS:

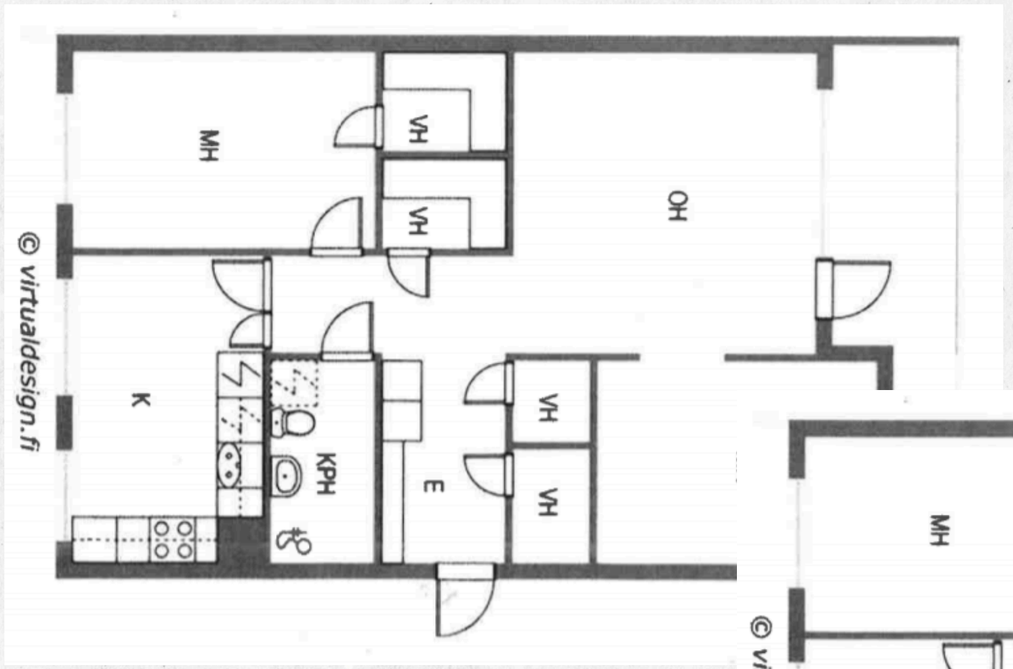
- Designed a model-guided generative pipeline for structural floor plan changes
- generated a labeled before / after dataset
- Trained and evaluated a change-type classification model
- Achieved understanding and detection of meaningful content changes.

PROJECT NOVELTY:

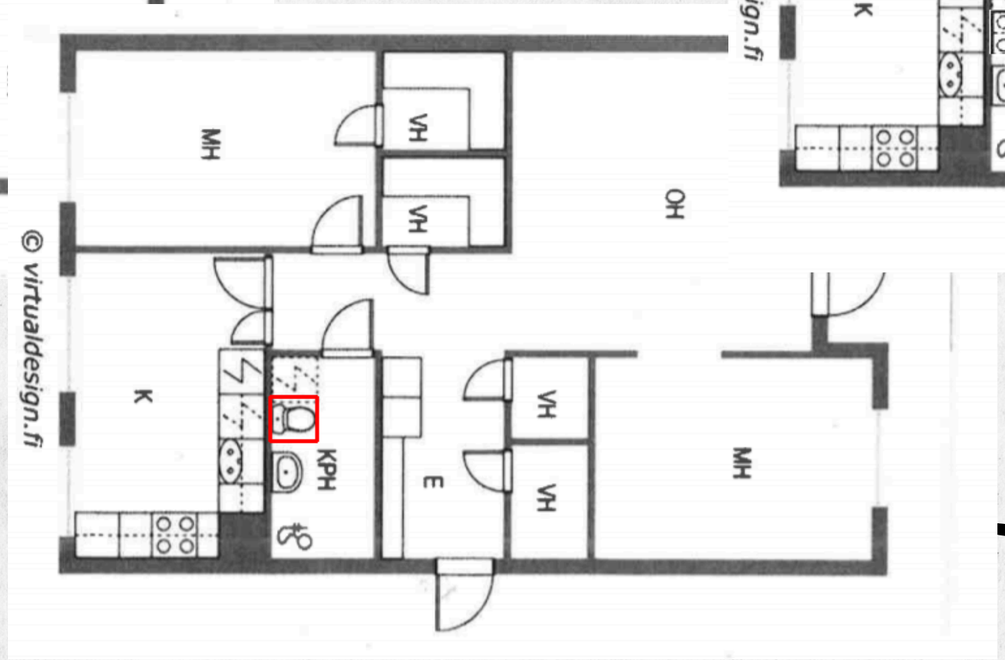
- No existing dataset for floor plan change detection
- Custom synthetic dataset created from real plans
- Controlled, explainable generation with precisely defined change annotations.

Toilet to Stove Example:

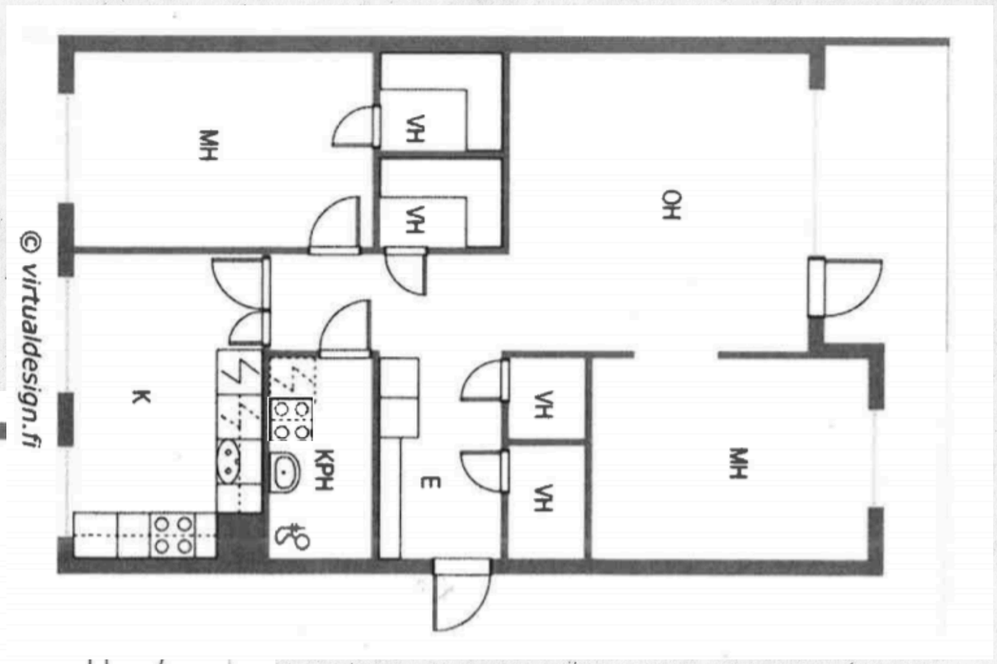
BEFORE:



BOX:



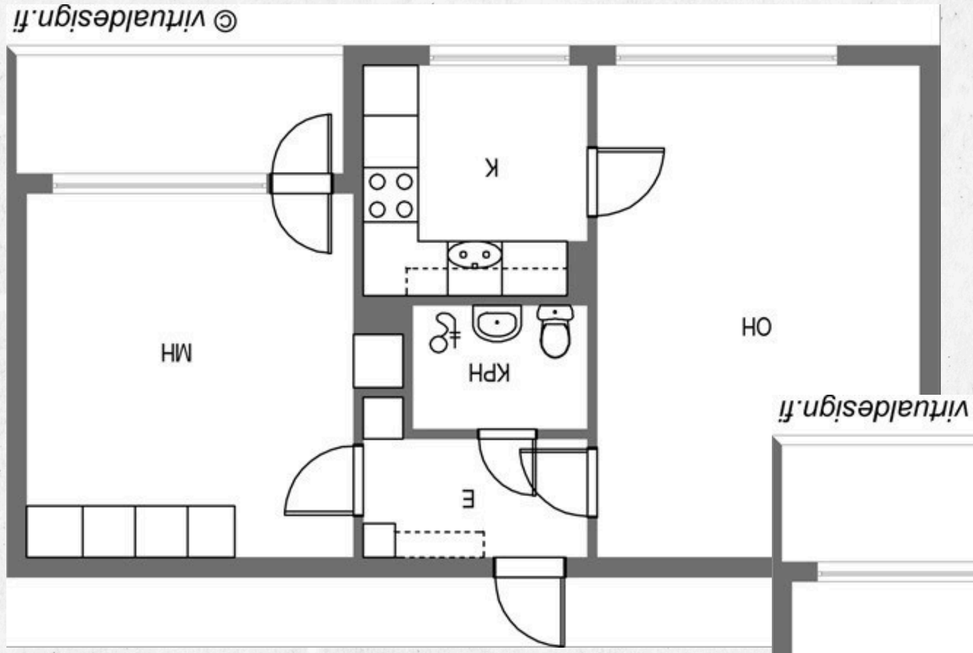
AFTER:



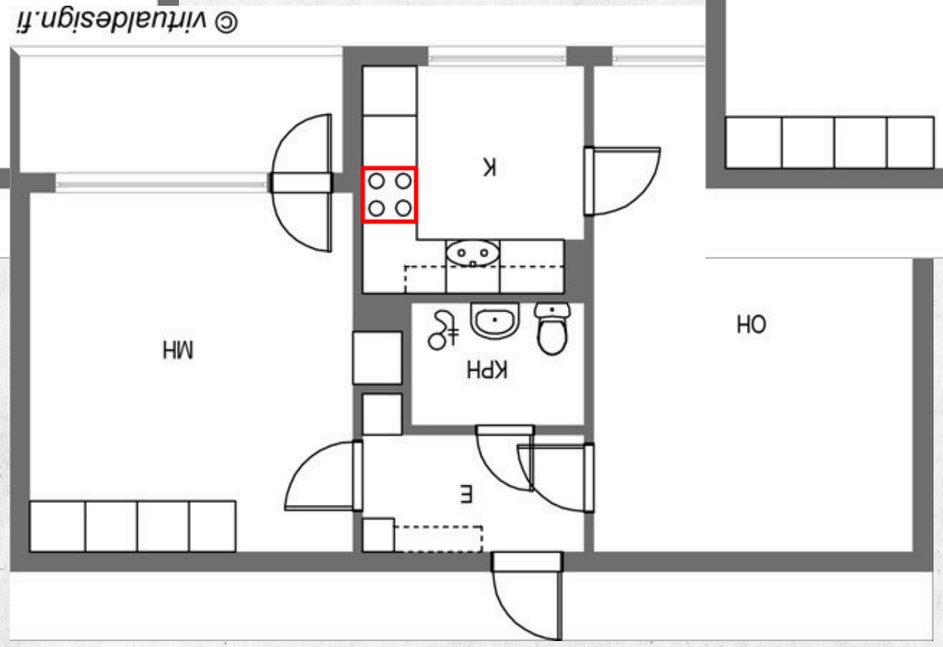
SUBSTANTIAL CHANGE CREATION

Erase Stove Example:

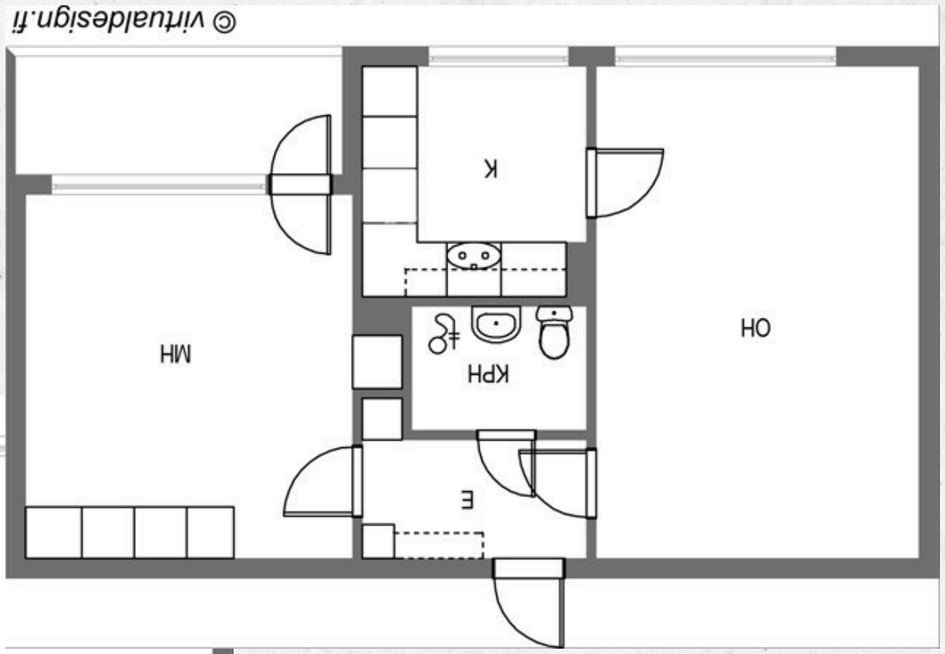
BEFORE:



BOX:



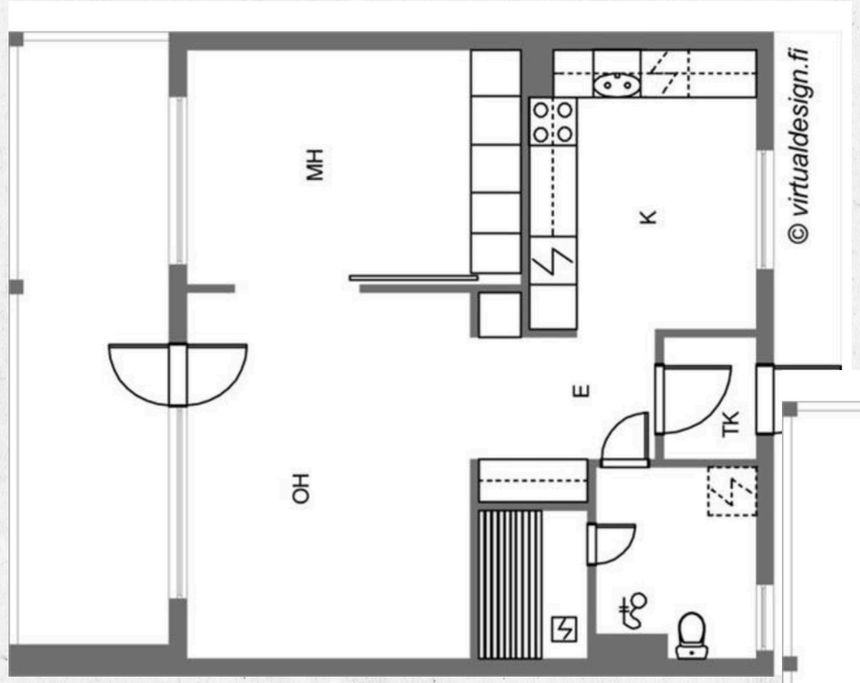
AFTER:



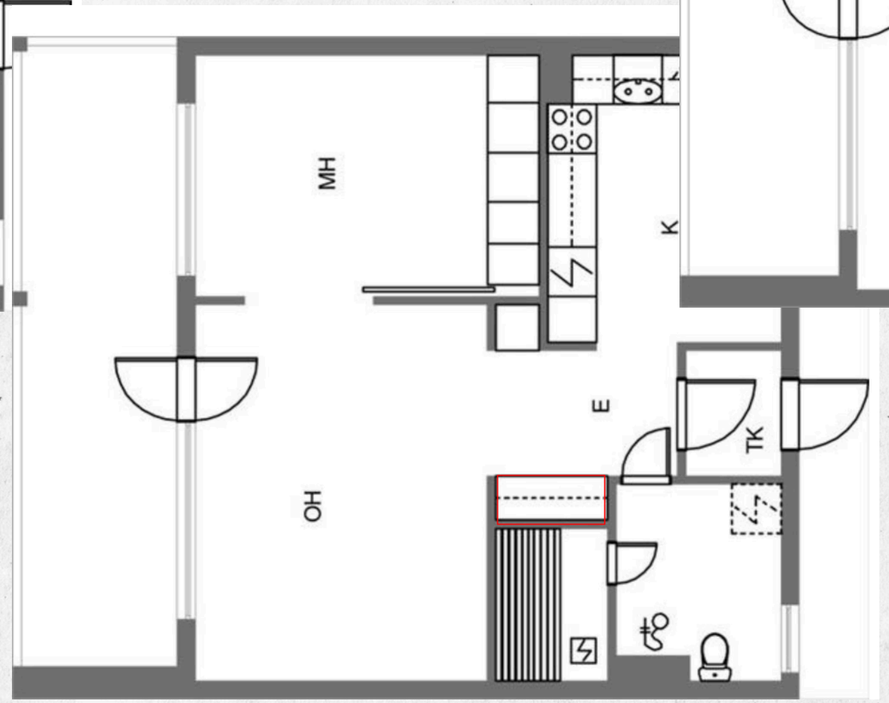
SUBSTANTIAL CHANGE CREATION

Erase Closet Example:

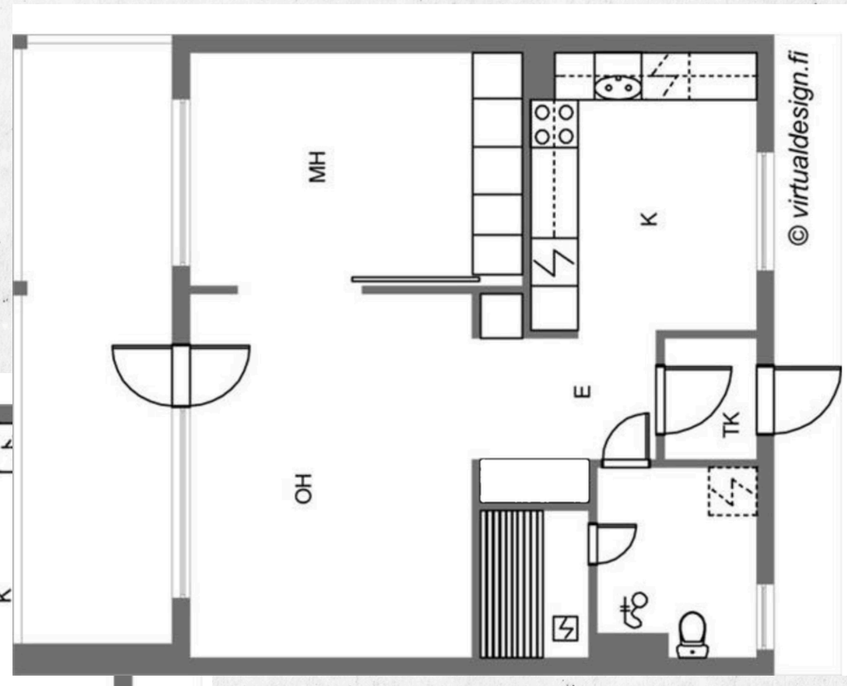
BEFORE:



BOX:



AFTER:



NON - SUBSTANTIAL CHANGE CREATION

Synthetic Data Generation Workflow

HUGGING FACE
Floorplans
Raw real-world images

→

DATA
cleaning & deduplication
Noise and duplicate removal

→

MANUAL ANNOTATION
(CVAT)
Object-level ground truth (455 images)

↓

DETR
Training
object detection & localization

←

SAM
segmentation
Precise object masking

←

INPAINTING
(Controlled)
Rule-based patch edits

↓

CHANGE-TYPE
Generation
significant & non-significant changes applied

→

MULTIPLE AFTER IMAGES
Creation
Up to 9 variants per source

→

SYNTHETIC DATASET
(Labeled Pairs)
Ready for classification

Synthetic Data Change Categories

SIGNIFICANT CHANGES:

Functional / structural modifications

- Remove stove
- Remove toilet
- Remove sink
- Replace sink with stove
- Replace toilet with stove

Impact core functionality

NON - SIGNIFICANT CHANGES:

Layout-level modifications

- Remove closet
- Add random closet
- Remove one-sided door
- Remove two-sided door

Limited functional impact

All change types and labels are defined during the controlled inpainting stage, resulting in a fully labeled synthetic dataset ready for classification.

Change-Type Classification Workflow

BEFORE / AFTER
Image Pairs
Synthetic image pairs

→

DETR INFERENCE
(Both Images)
Detect objects independently

→

OBJECT SET
Comparison
Detect changed objects

↓

HUNGARIAN MATCHING
One-to-one object alignment

←

FEATURE VECTOR
Extraction
Encode object-level differences

↓

LOGISTIC REGRESSION
Training
Learn change-type decision

→

THRESHOLD
Selection
Chosen by highest F1-score

→

FINAL PREDICTION
Predicted change type

Methodology Summary

SYNTHETIC DATA GENERATION:

- DETR for object detection and localization
- SAM for precise region segmentation
- Controlled patch-based edits to create before / after pairs

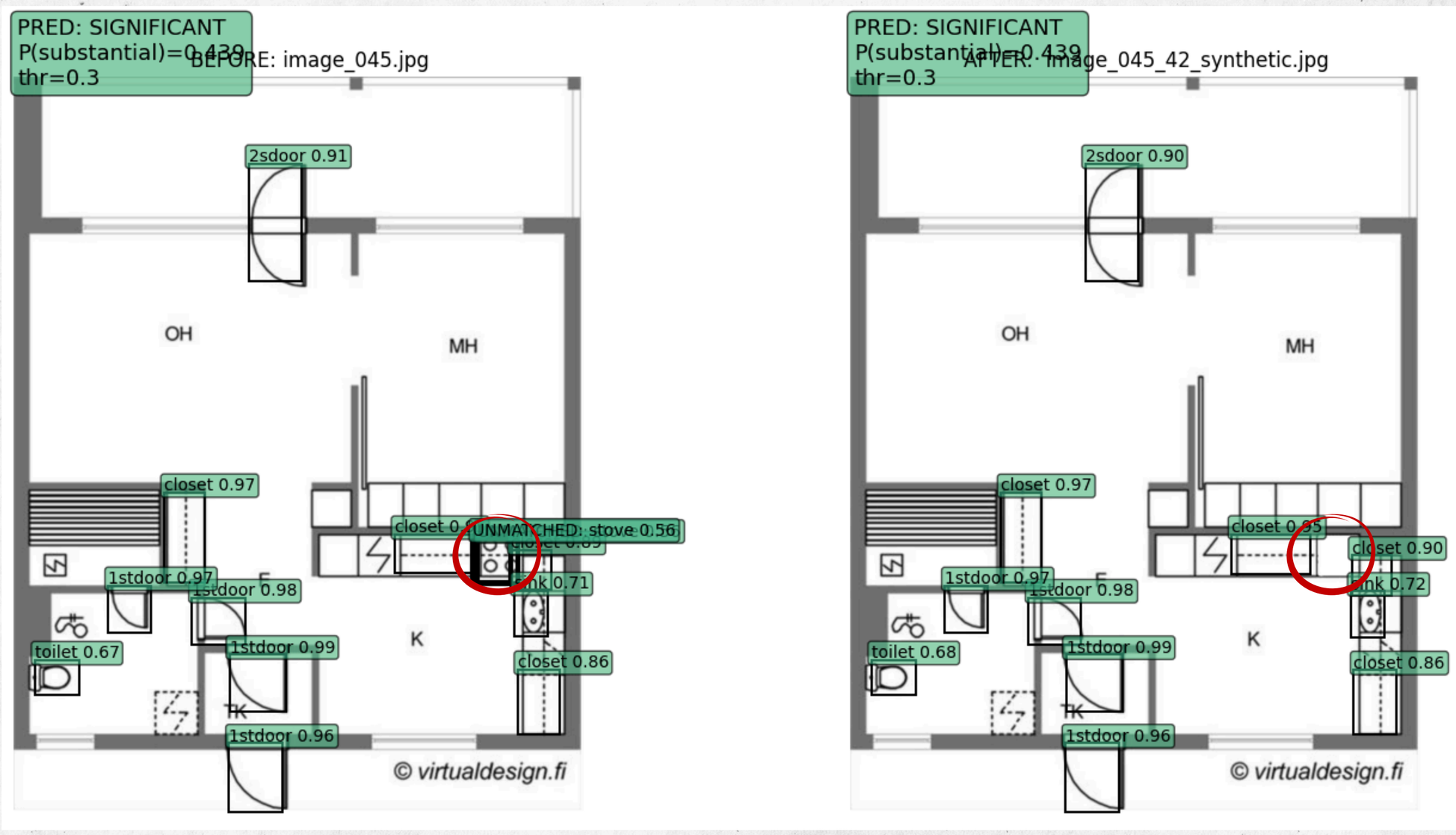
EVALUATION STRATEGY:

- Accuracy, Precision, Recall, F1-score
- Threshold sweep on prediction confidence
- Threshold = 0.3 selected for best performance on meaningful changes

CHANGE-TYPE CLASSIFICATION:

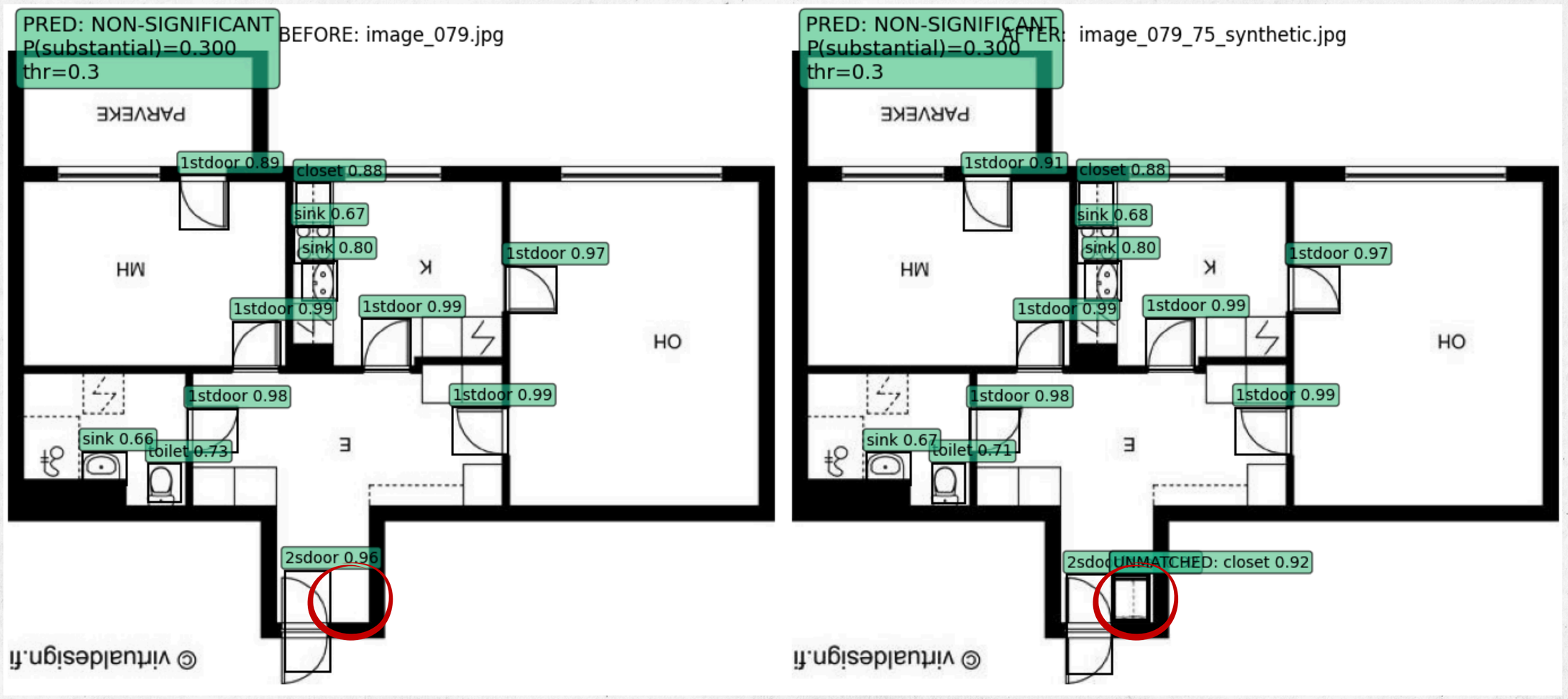
- DETR-based classifier with Hungarian Matching
- Feature Engineering produced by the Hungarian Matching
- Logistic Regression model
- Output: predicted change type between image pairs

predicted: significant ✓



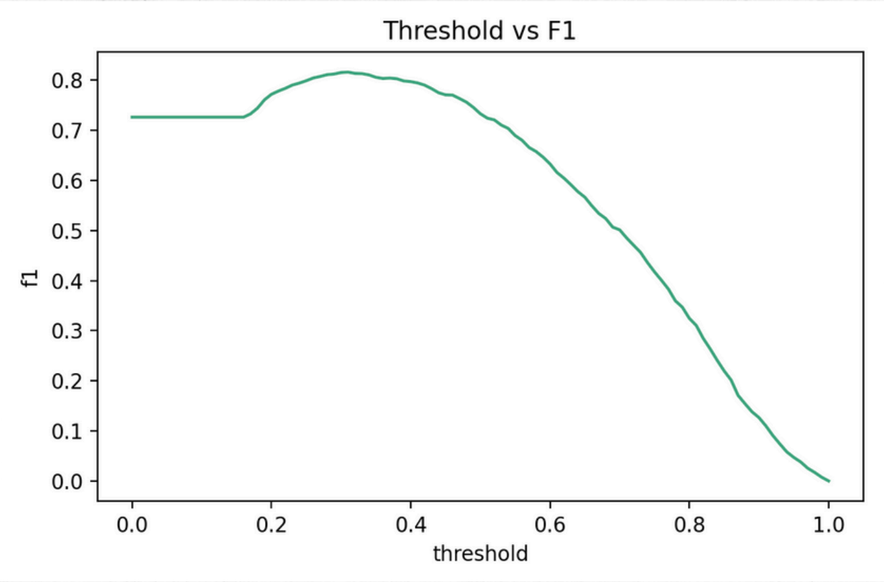
REMOVED STOVE

predicted: non-significant ✓

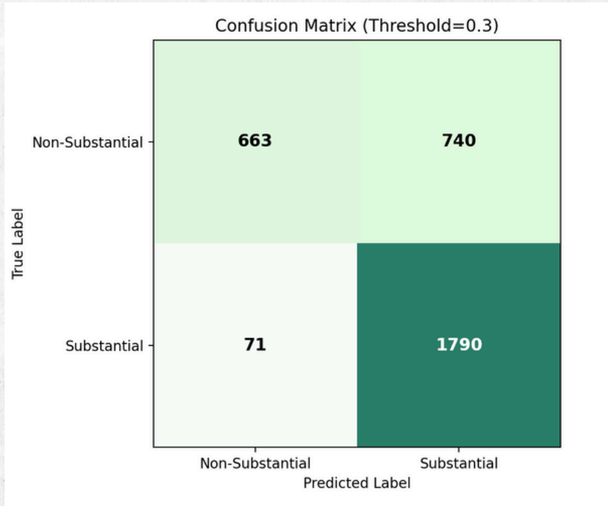


ADDED CLOSET

Results



THRESHOLD = 0.3 SELECTED FOR F1 SCORE OF 0.81



HELPS ANALYZE FALSE POSITIVES AND FALSE NEGATIVES.

HIGH-CONFIDENCE PREDICTIONS ARE RELIABLE

Conclusion

SUMMARY:

- Detected meaningful changes in floor plan images
- Created a novel synthetic dataset where none previously existed
- Built a large labeled before/after image set

FUTURE WORK:

- Expand and refine change types and scenario complexity
- Combine synthetic and real-world data
- Use confidence to flag uncertain cases

KEY INSIGHTS:

- Data quality and change definition are more critical than model complexity
- Different models perform better for different change types
- Threshold selection has a strong impact on practical performance