



D.1.2.2

Database of infrared images used for structural and damage training of ML/AI system



Italy – Croatia



Project acronym	STRENGTH
Project full title	STRategies for assessing climate change and natural hazards' impact on urban ecosystems, increasing resilience to ENVIRONMENTAL hazards, and promoting territorial GrowTH
Programme	Interreg Italy-Croatia 2021-2027
Start date	01/04/2024
End date	30/09/2026
Project ID	ITHR0200318

Deliverable Title	D.1.2.2 - Database of infrared images used for structural and damage training of ML/AI system
Activity	Digital monitoring to detect seismic vulnerability
WP	1
WP Leading Partner	FESB
Contributing Partners	FGAG, UNIFE
Dissemination level	Confidential <i>or</i> Public
Version	Finalised
Date	30/09/2025





Table of contents

Executive Summary.....	4
1. Database organization	5



Executive Summary

This document defines the **data architecture and file-management workflow** for the thermographic façade dataset used in the project “*Survey of Deep Learning Architectures for Façade Layout and Masonry Material Segmentation.*” It provides a unified standard for how data is stored locally, how reduced datasets are exported for project partners, and how overall access to the dataset is maintained during the project lifecycle.

Purpose

The structure serves two main goals:

1. **Local neural-network training:** full-resolution raw data remains on the local machine for efficient processing.
2. **Cloud partner access:** due to large file sizes, only reduced or selected exports are synchronized to a shared cloud/drive location for collaborators.

Data Synchronization Strategy

- **Local Machine:** holds the complete dataset, including all high-resolution thermographic images.
- **Cloud/Drive:** contains a lightweight, reduced version of the dataset intended for external partners.
- Only **partial files** (specific image sets) are uploaded due to storage and bandwidth limitations.

File Types

- Images stored as **.png** or **.jpg**.
- Metadata and logging stored in per-location **README** files.

Top-Level Organization

The dataset is organized under a central **Project File** directory. Each survey location (e.g., *Kaštela 1*, *Kaštela 2*, *Kaštela 3*) contains:

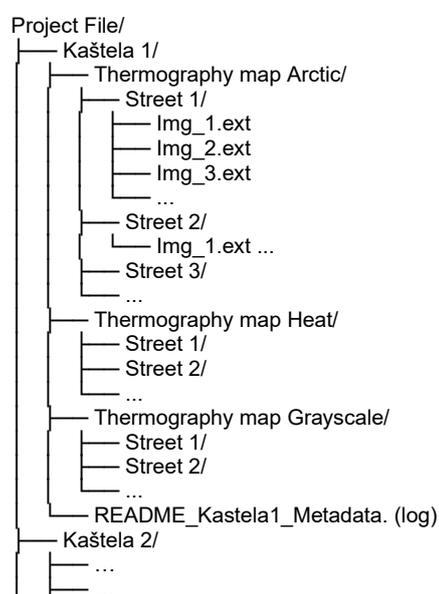
- Multiple **thermography map modalities:**
 - *Arctic*
 - *Heat*
 - *Grayscale*
- Each modality contains subfolders by **street** (Street 1, Street 2, ...).
- Each street folder contains sequentially named thermographic images (*Img_1.ext*, *Img_2.ext*, ...).
- Each location includes a metadata log (**README_KastelaX_Metadata**).



1. Database organization

For the purposes of training networks mentioned in Survey of Deep Learning Architectures for Façade Layout and Masonry Material Segmentation. This document serves as a baseline on how the database is populated and what is exported for partners during the project because of the large file sizes.

File Structure:



Project File/ Top-level container for the whole dataset.

Files are synced between:

- Local machine (full raw dataset for NN training)

- Cloud/Drive (reduced/export version for partners if/when needed)

.ext – png/jpg

Only partial files will be exported and synced with cloud version because of large files sizes.

For the duration of the project files can be accessed on link:

[STRENGTH](#)

