Imperial College London

Autonomous Intelligent Drones

Christos-Savvas Bouganis
Christos-savvas.bouganis@imperial.ac.uk

intelligent Digital Systems Lab
Dept. of Electrical and Electronic Engineering

www.imperial.ac.uk/idsl

The team

Stylianos I. Venieris
Machine Learning

Alexandros Kouris
Machine Learning, Robotics

Konstantinos Boikos
Computer Vision, SLAM

Manolis Vasileiadis
Computer Vision

Mudhar Bin Rabieah
Machine Learning

Nur Ahmadi
Brain-Machine Interface

Christos-Savvas Bouganis
iDSL Lab Director
Imperial College London

Vision: Autonomous Intelligent Drones

Goals:
- Perceive the environment
- Understand the environment
- Interact with the environment

Challenges:
- Low latency
- Low power
- Adaptation

Conventional and Unconventional Embedded Platforms for Compute

GPUs – Tegra K1, X1 and X2
DSPs – Qualcomm Hexagon,
Apple Neural Engine, …



✓ High throughput
✗ Low latency
✗ Low power
✓ Tools

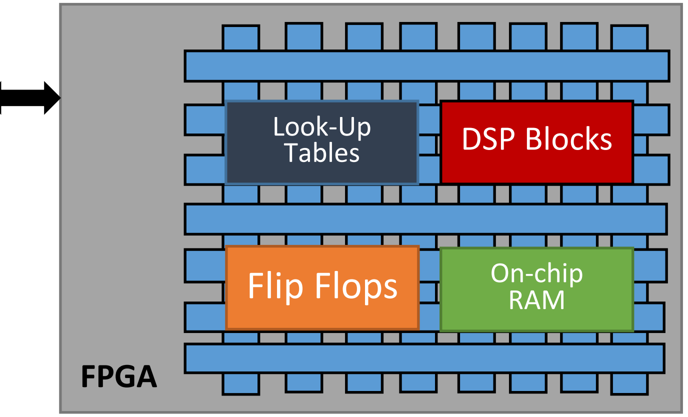
customisation

Ambarella
Myriad X

FPGAs

- Custom datapath
- Custom memory subsystem
- Programmable interconnections
- Reconfigurability

External Memory (DRAM)



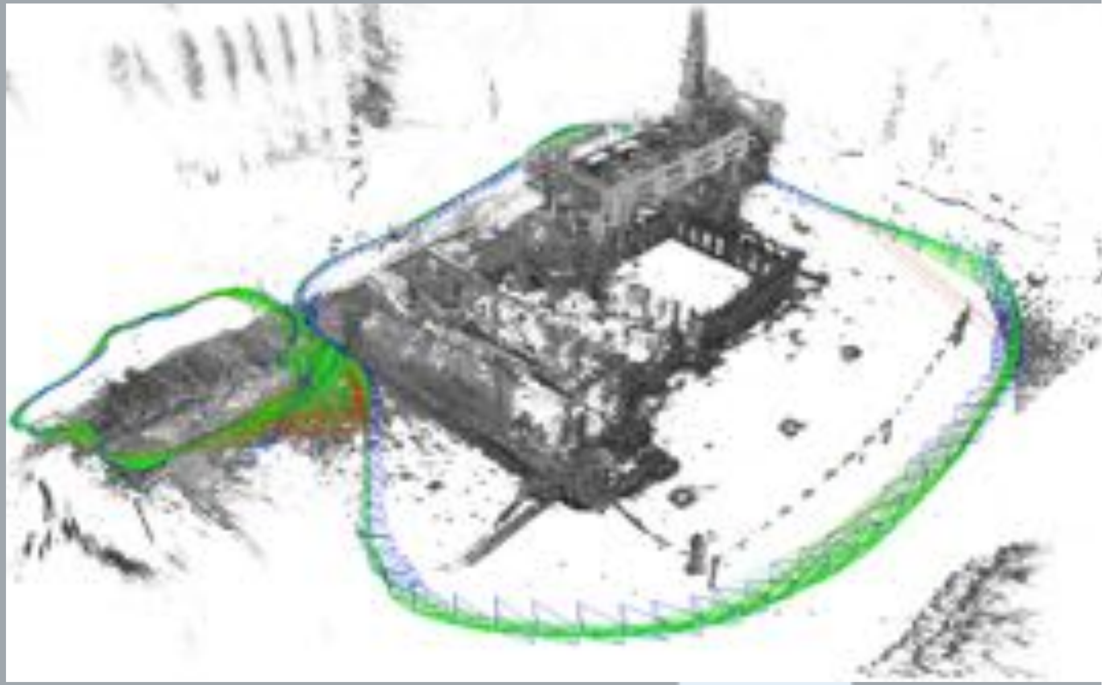
✓ High throughput
✓ Low latency
✓ Low power
✗ Tools

Challenge: Huge design space
Our Approach: Automated toolflows

Research Areas / Challenges

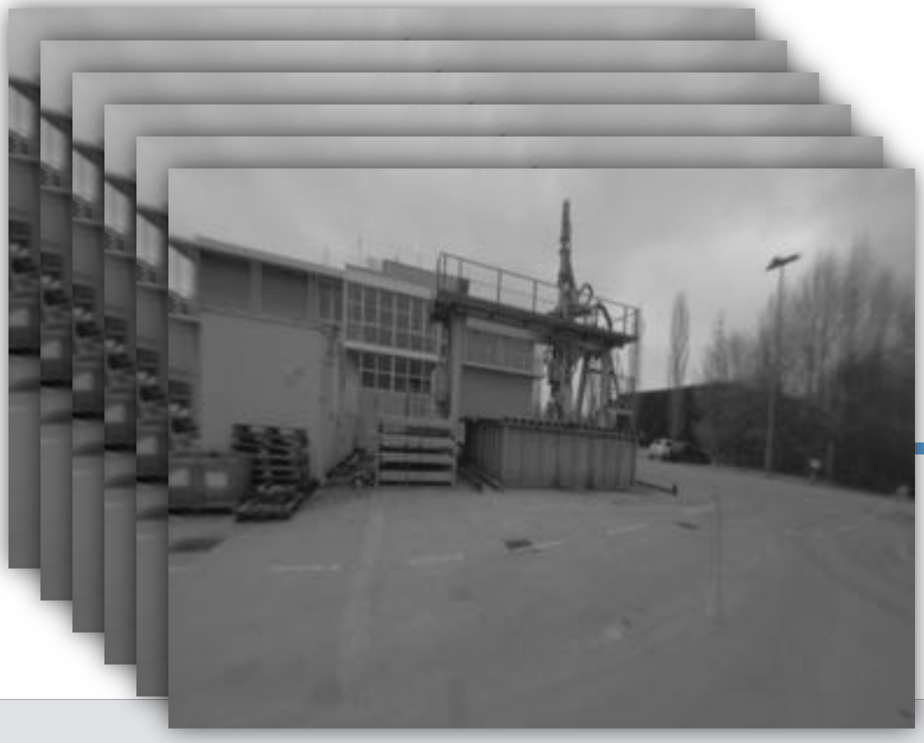
- SLAMSoC
- Learn to Fly

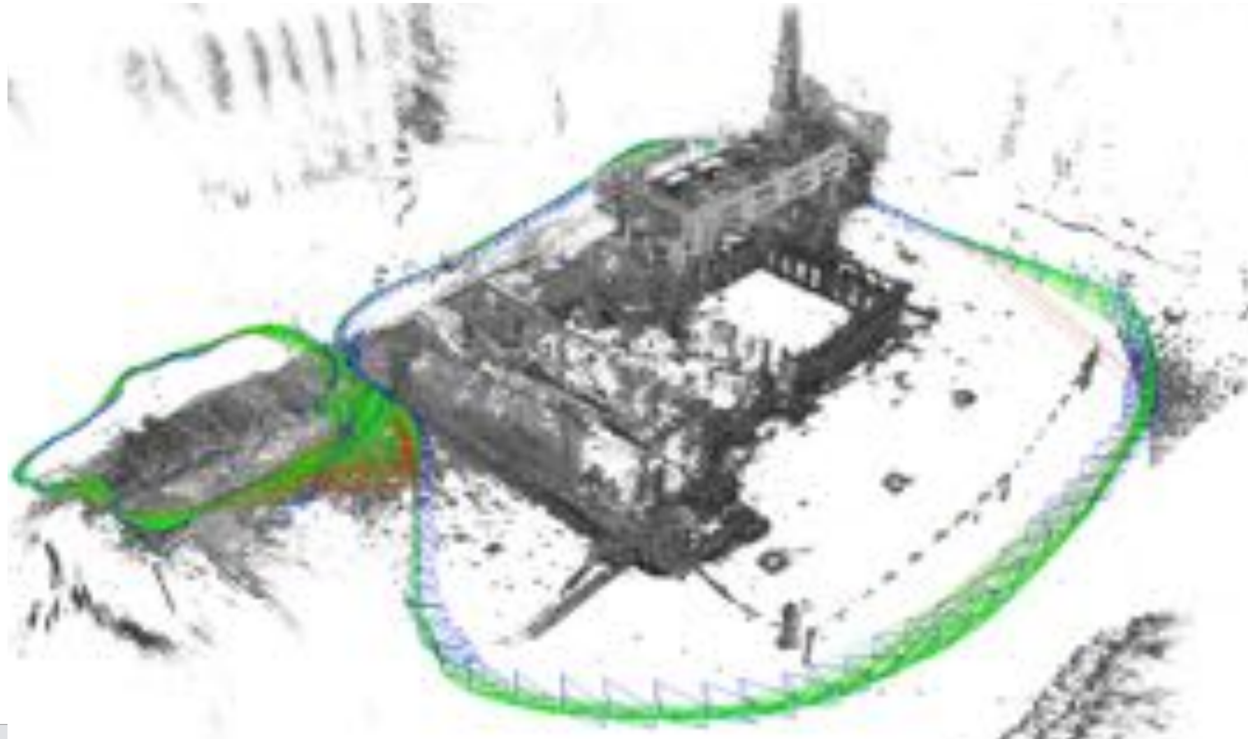
Topic #1: SLAMSoC



Use a series of observations to simultaneously perform Localisation and Mapping

- Tracking (Localisation): Online pose estimation of the sensor and robot.
- Mapping: Fuse observations into a coherent model of the environment

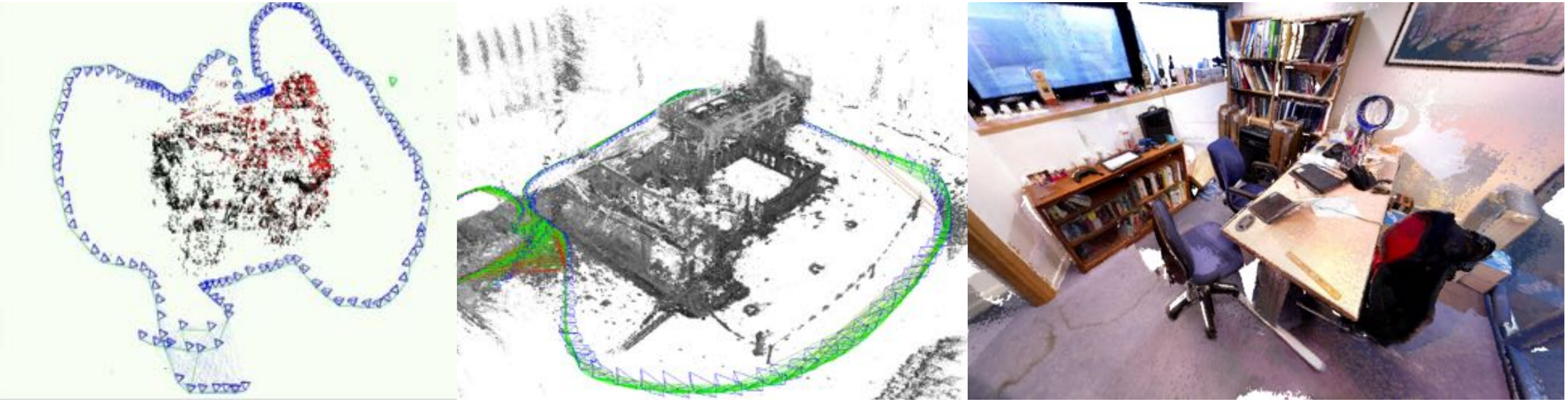




Source: LSD-SLAM (J. Engel et al.)

Challenges in Embedded SLAM

- Emerging algorithms have high complexity and bandwidth requirements
- Field still in a state of constant change
- **Tracking** robustly needs high framerate and low latency



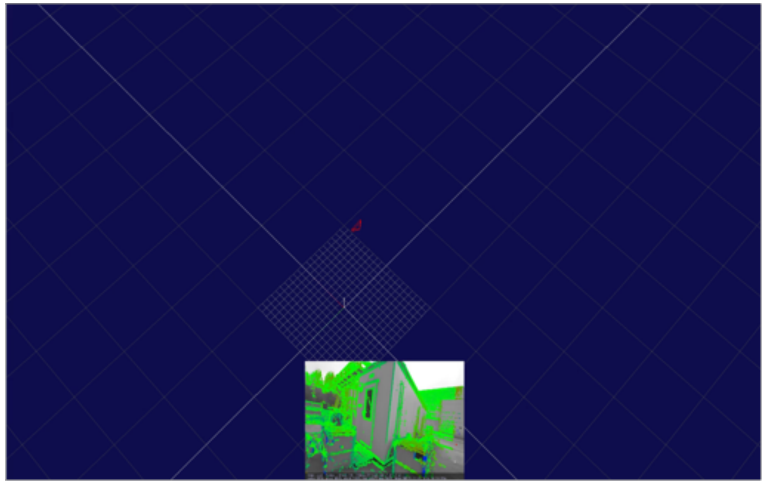
Sparse
Mobile CPU

Semi-Dense
High-end Desktop

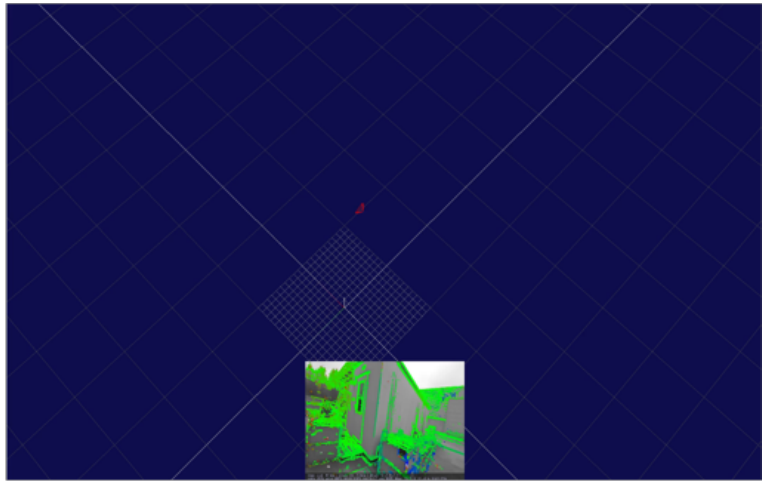
Dense SLAM
High Performance GPU Acceleration

Sources: ORB SLAM (R. Mur-Artal), LSD-SLAM (J. Engels et al.), ElasticFusion (T. Whelan et al.)

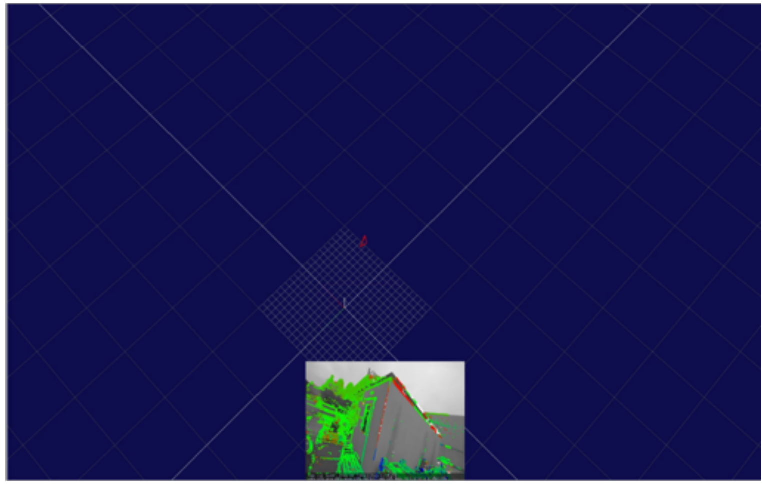
Importance of High Performance Tracking



- ➢ Camera rate (30fps)
- ➢ Intel i7-4770



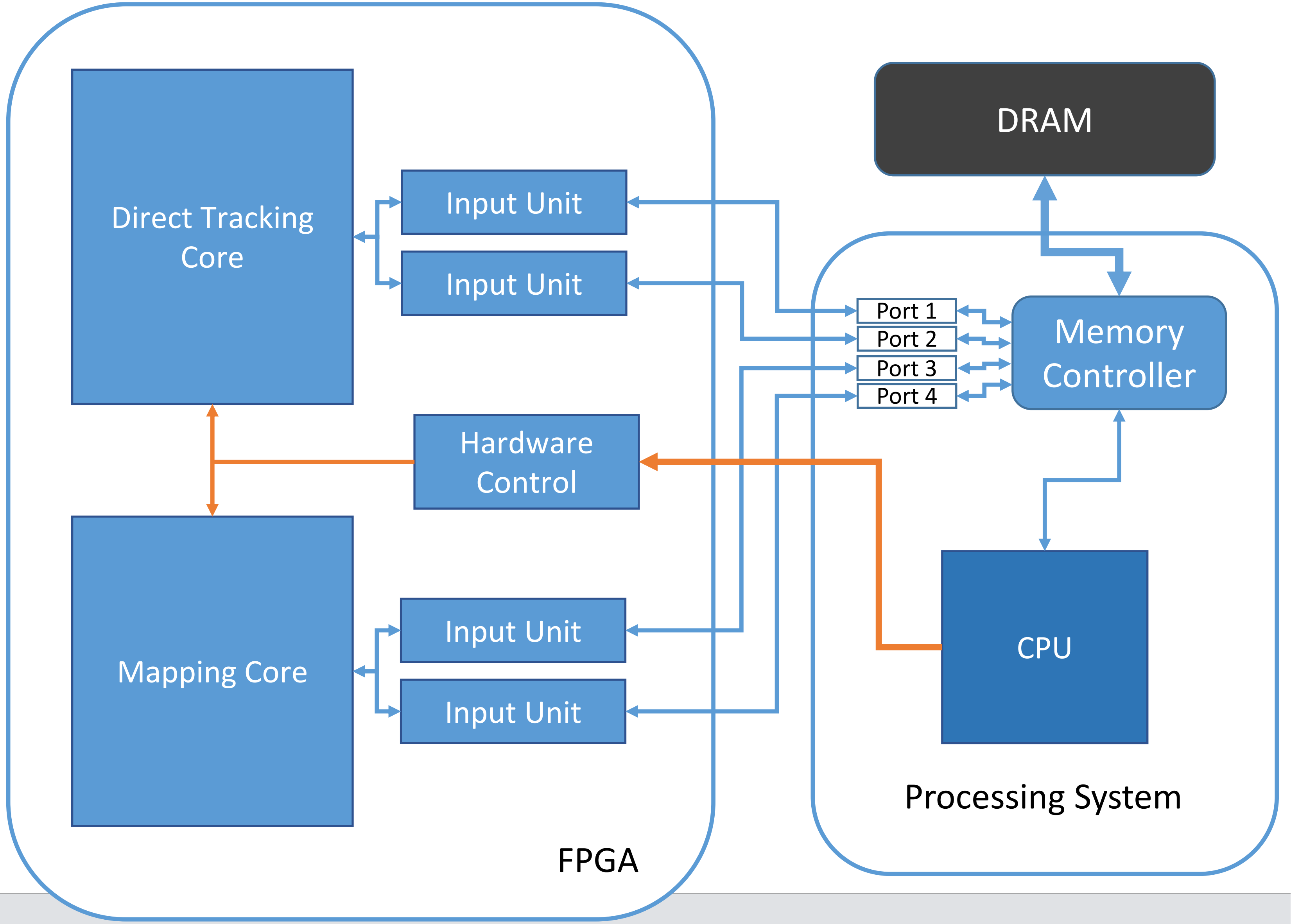
- ➢ Drop frames (15 fps processed)
- ➢ Position Drift, Error accum.



- ➢ Processing <8 frames/s
- ➢ Lost tracking

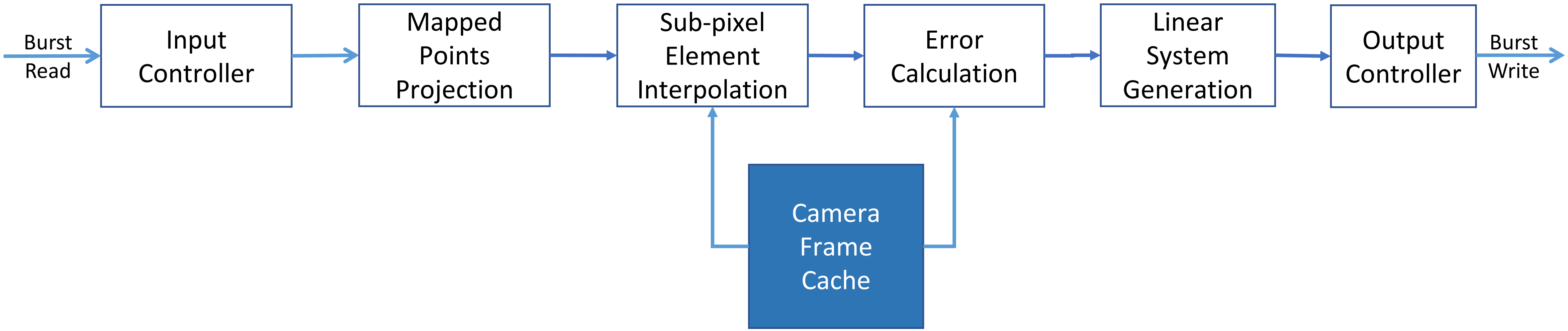
Proposed system

- Common memory space and Direct Memory Access
- Hardware high-level control from CPU
- Both operate simultaneously
- Buffered high memory bandwidth



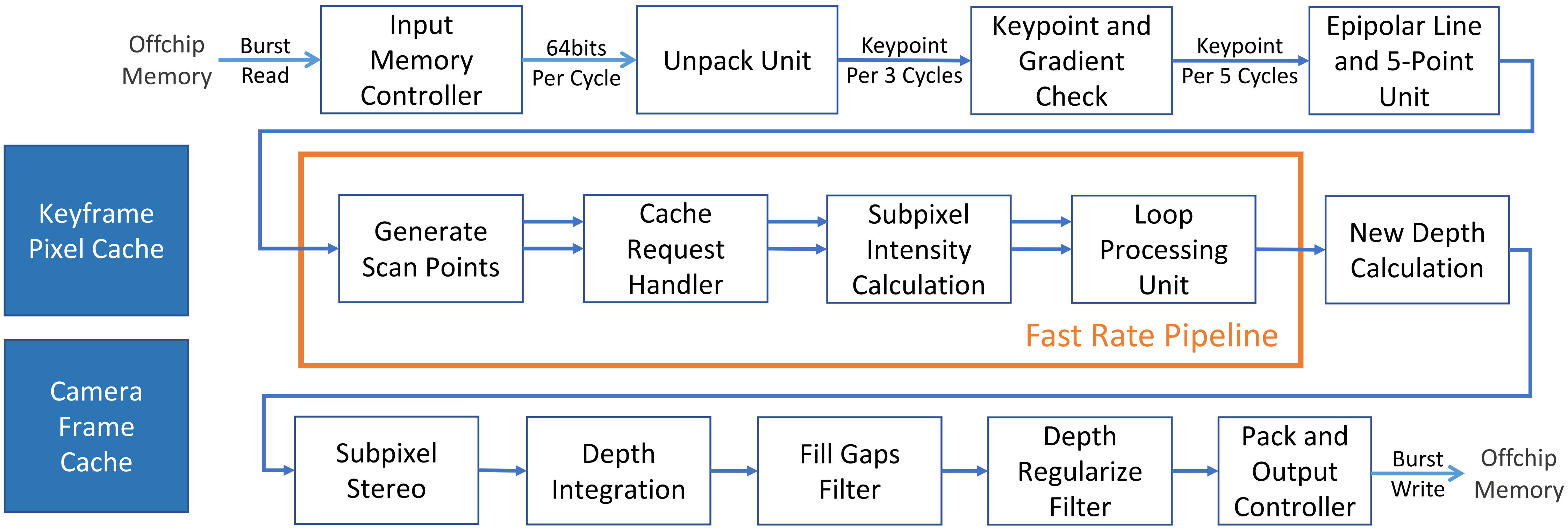
Direct Tracking Core

- Streaming Dataflow – Designed with High Level Synthesis
- Splitting computation into smaller blocks allows better optimisation
- Separating control flow from computation leads to a better design
- Redundant computation proved more efficient than going to memory

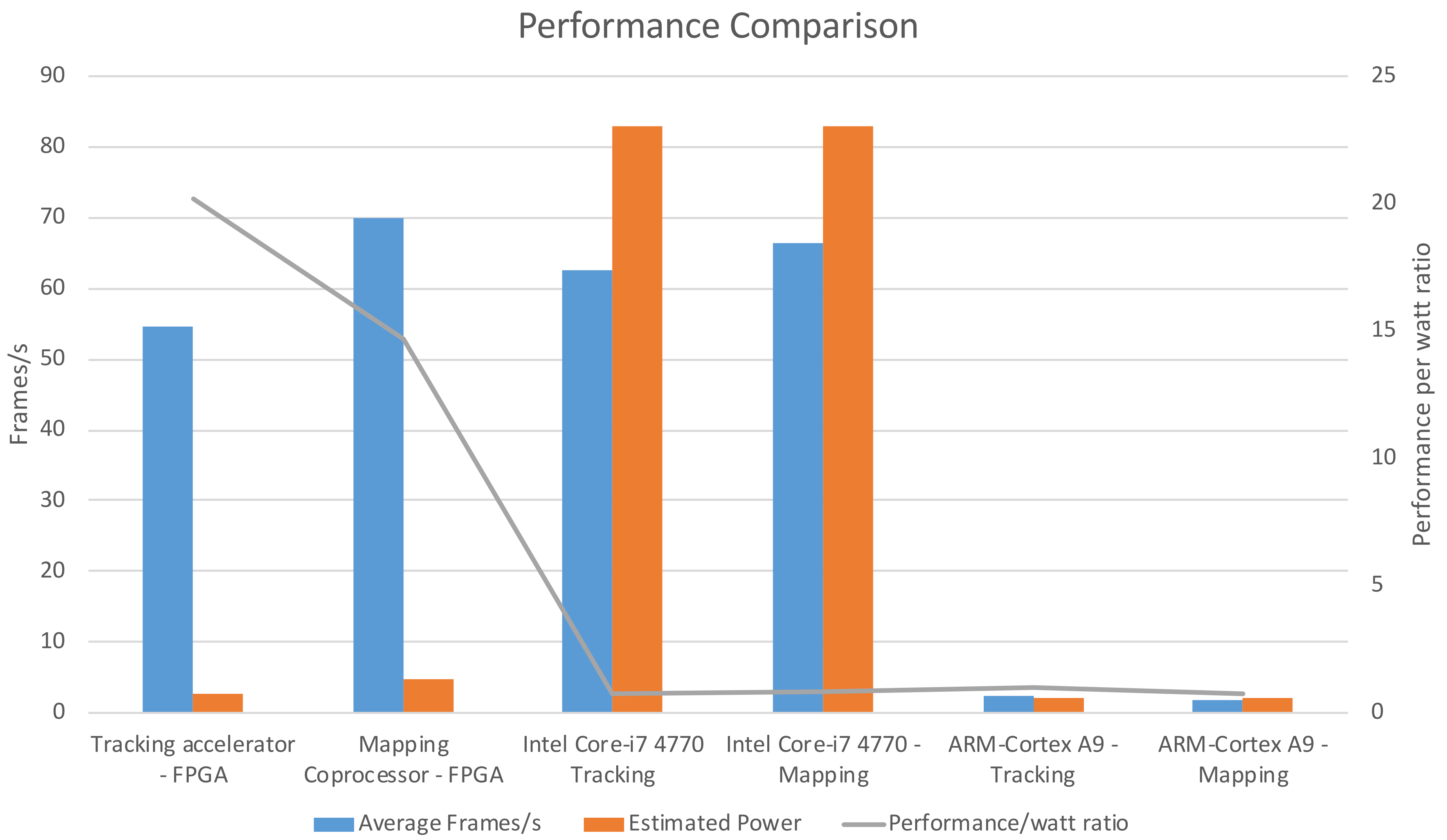


Mapping Coprocessor

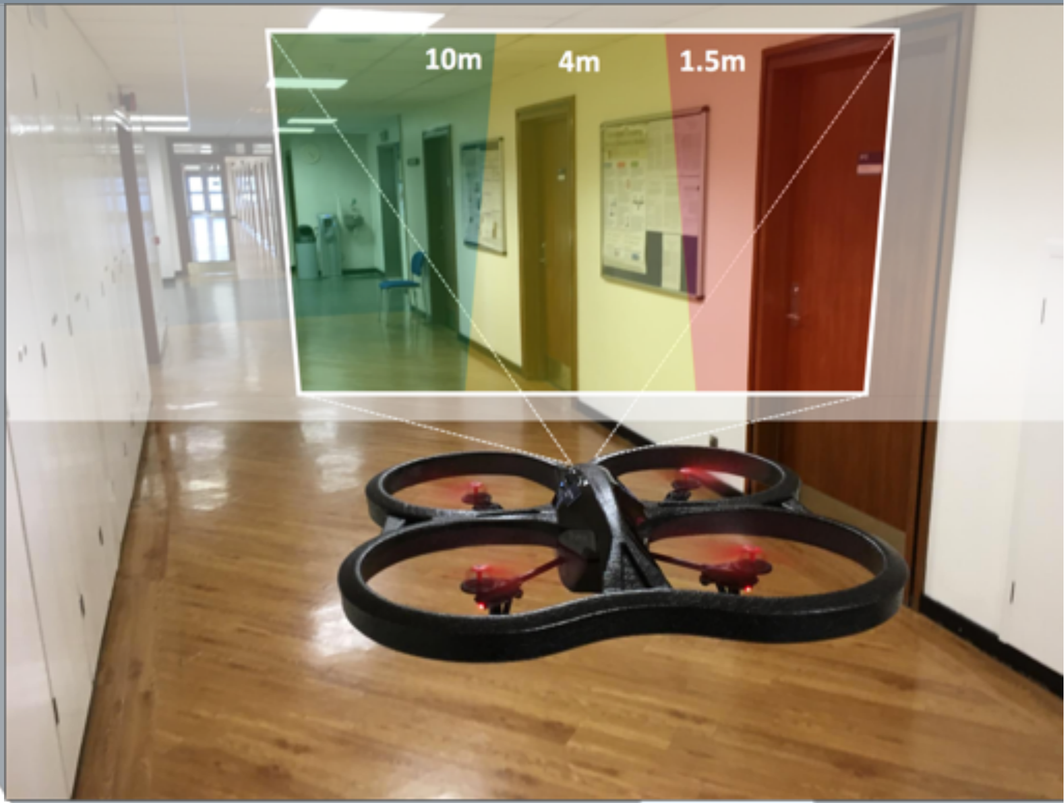
- Variable rate pipelines
- Streaming dataflow processing, combined with local caches for random-access patterns



Comparison with other platforms

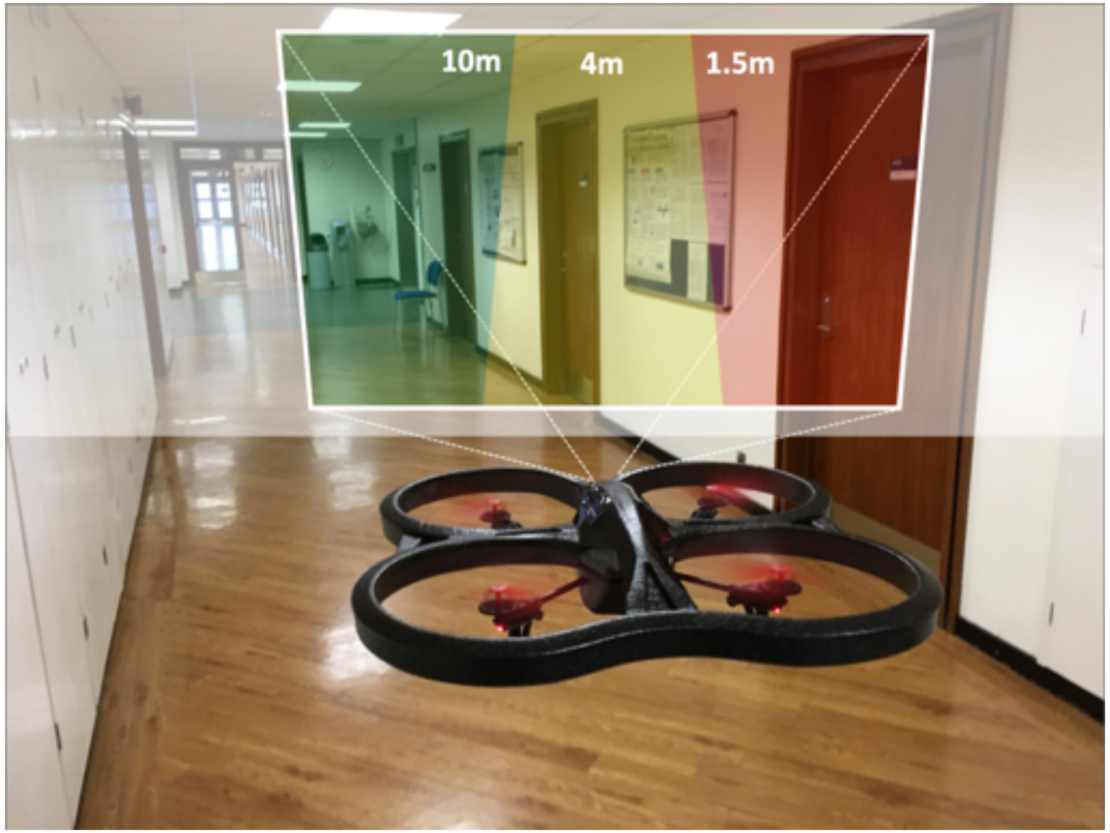


Topic #2: Learn to Fly



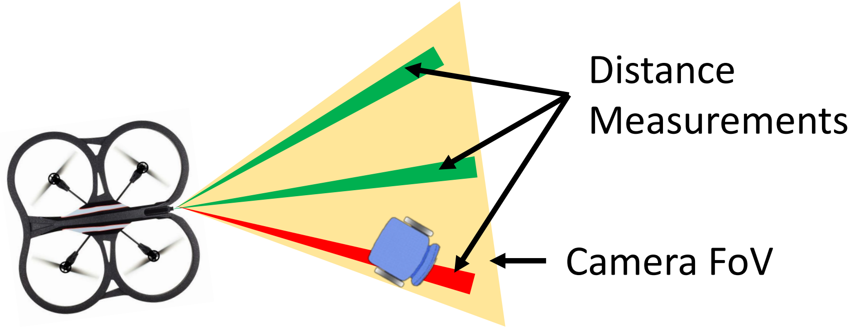
Self-Supervised approach for autonomous navigation:

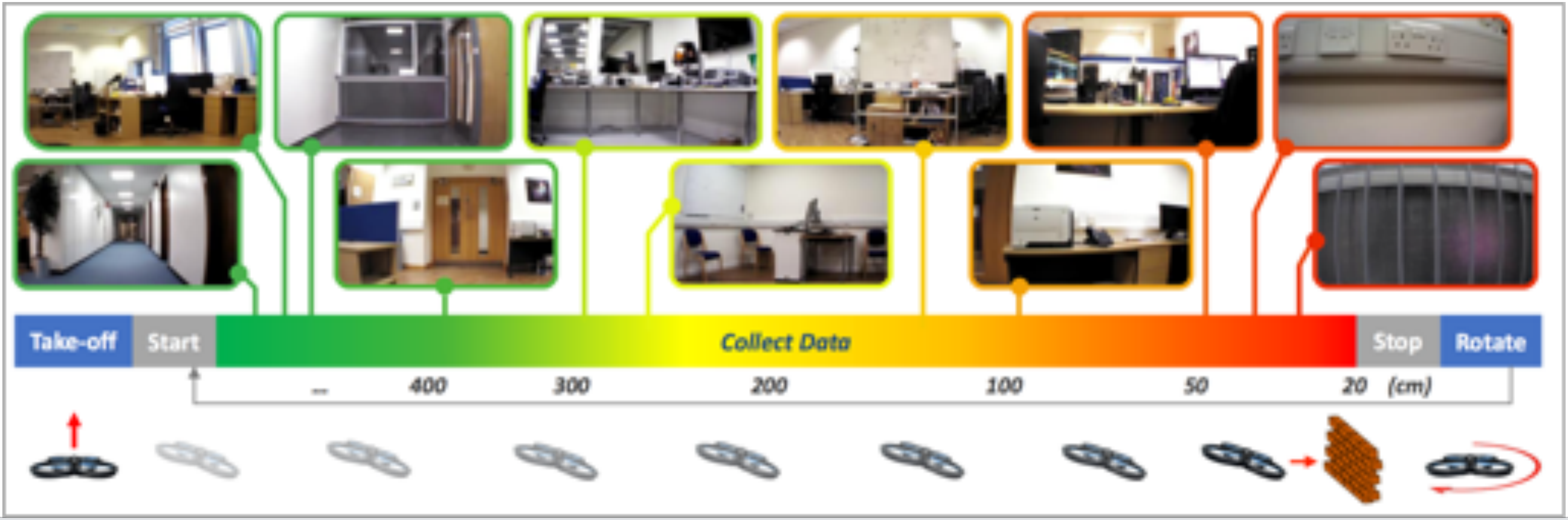
- Exploits solely on-board camera's visual input
- Regression CNN to predict distance-to-collision
- Local path planner to modulate velocities



Self-Supervised Data Collection & Annotation

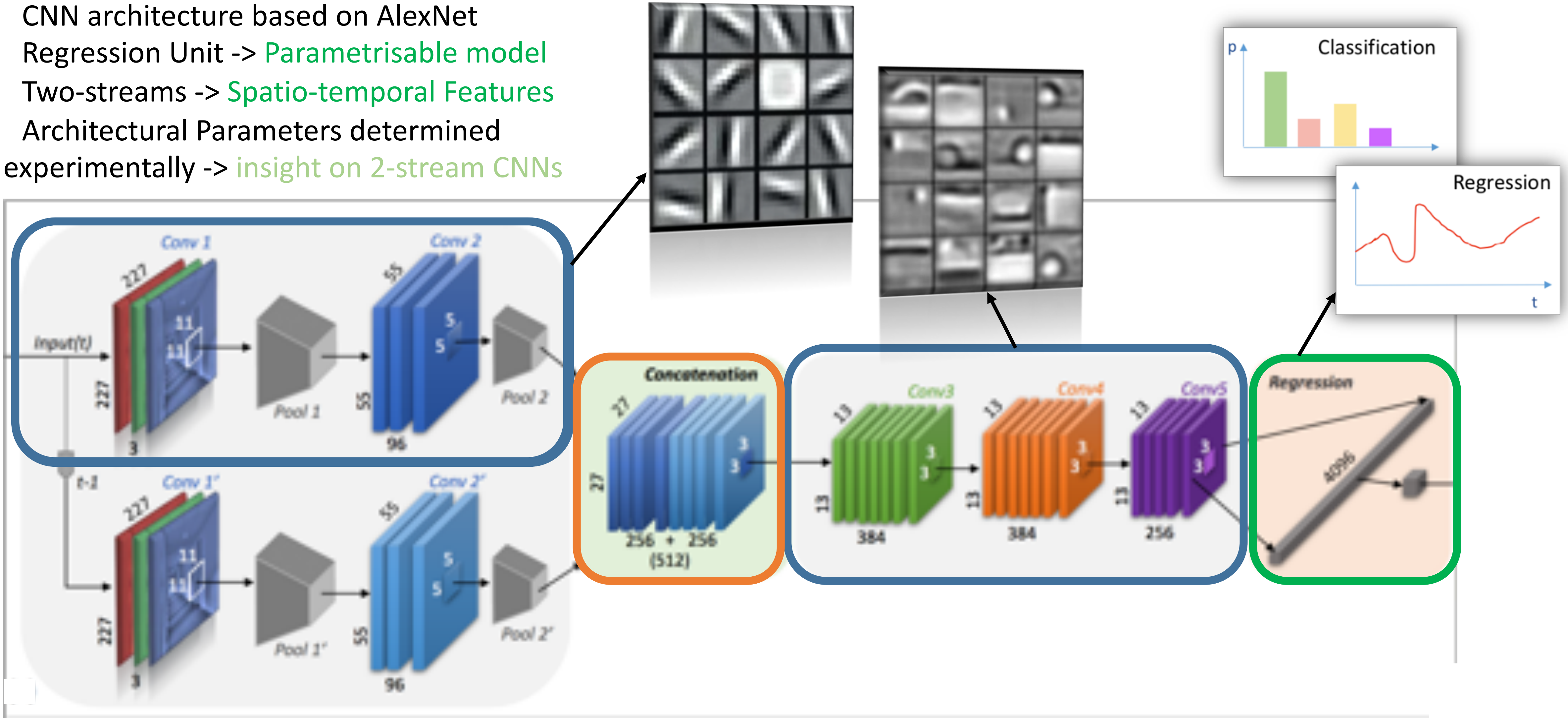
- Robust Deep Learning models require tons of data
- External Distance Sensors to automate the collection
- Indoor Flight Dataset:
  - Annotated with real-distance values
  - 300.000 samples in 2000 trajectories





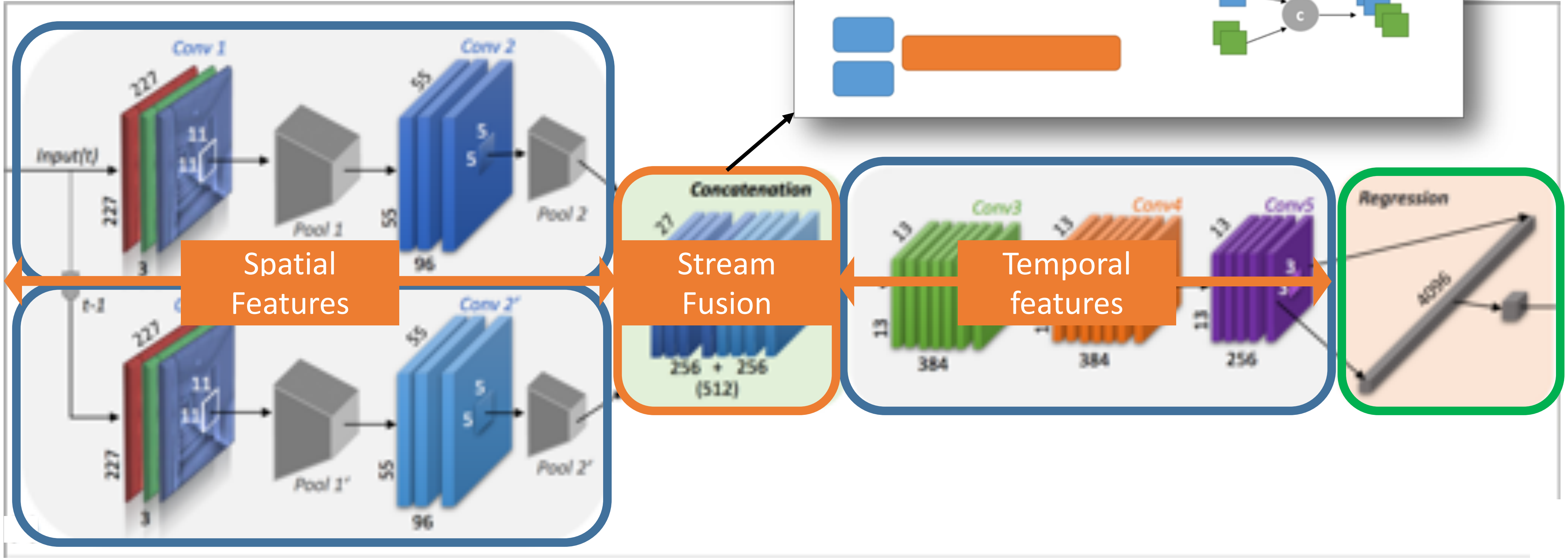
2-stream Regression CNN Architecture & Training

- CNN architecture based on AlexNet
- Regression Unit -> Parametrisable model
- Two-streams -> Spatio-temporal Features
- Architectural Parameters determined experimentally -> insight on 2-stream CNNs



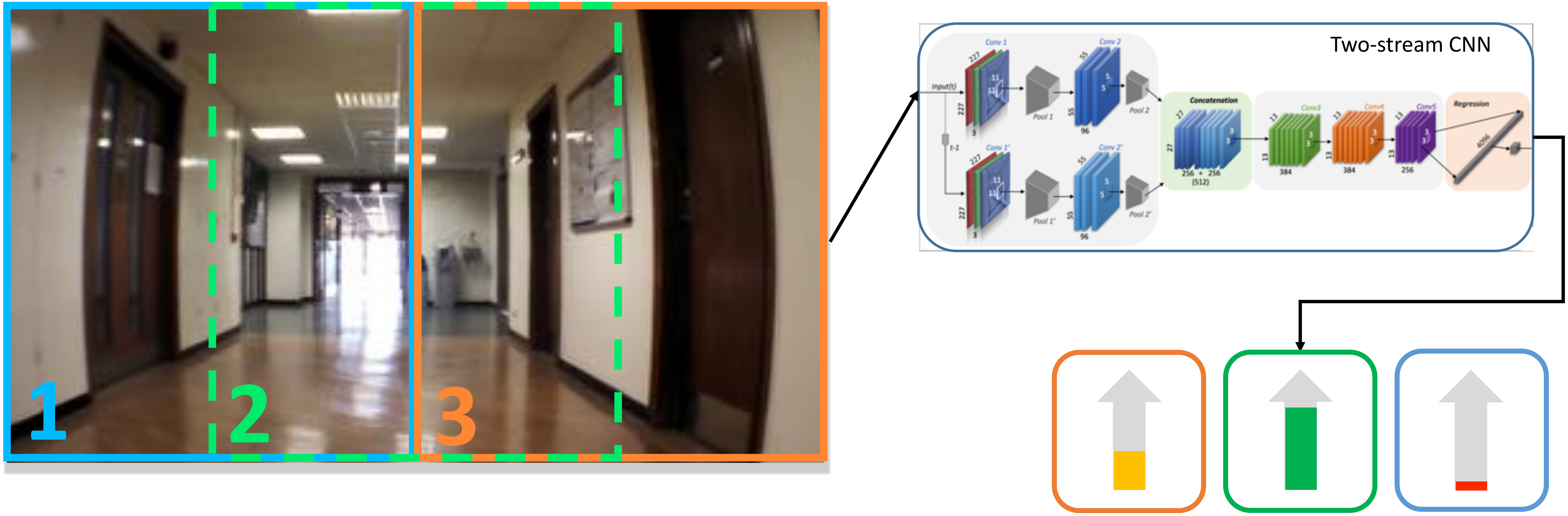
2-stream Regression CNN Architecture & Training

- CNN architecture based on AlexNet
- Regression Unit -> Parametrisable model
- Two-streams -> Spatio-temporal Features
- Architectural Parameters determined experimentally -> insight on 2-stream CNNs



2-stream Regression CNN Architecture & Training

- The CNN predicts the distance-to-collision for three partly overlapping windows of the image



Distance-to-collision Predictions

Quantitative Evaluation: End-to-end UAV Navigation Task

Classification vs Regression-based approach