ThermalNet

Batuhan Yıldız, Can Mızraklı, Gökay Toğa
Advisor: Prof. Dr. Tolga Kurtuluş Çapın

TED University

Problem Statement

- Paired RGB-thermal cameras mounted on drones, rovers or gimbals often drift out of alignment due to vibration, zoom and pan/tilt motion. Mis-registration degrades fused analytics and operator situational awareness. **ThermalNet** learns to **extract, match and geometrically align** cross-modal key points so that both streams remain pixel-aligned in real time, whatever the movement.

System Overview

- Input: synchronous RGB–thermal frame pair (640 × 480).
- Keypoint Extraction: SuperPoint-style CNN produces {k, d, s} per modality.
- Matching: 18-layer SuperGlue GNN computes feature matches and confidence masks.
- Homography Estimation: differentiable DLT + Smooth-L1, reprojection and sub-pixel consistency losses, robust to NaN/Inf.
- Sub-Pixel Refinement: RAFT optical flow refines matches to <0.5 px accuracy, thus quantifying the movement of the matches.
- Scale aware loss applies bilinear interpolation and refines the matches by checking them in different scales.
- Output: 3 × 3 homography H that warps the thermal frame onto the RGB view; optional overlay of inlier lines.

Data & Training

- Dataset: in-house ThermalHomographyDataset, ≈ 640 × 480 RGB/thermal pairs + ground-truth H; graceful fallback to identity when missing.
- Augmentation: colour-jitter, flips, rotations, Gaussian noise & cosine pseudo-thermal synthesis.
- Optimiser: Adam, LR = 1 × 10-4, 100 epochs, mixed-precision AMP; training on A100.

Results

- After 100 epochs the network converged to stable loss and produced visually consistent alignments in both day and night scenes; green lines denote inlier matches linking RGB→thermal features.
- All functional/robustness tests (grayscale conversion, NaN safety, inference tensor shapes, overfitting sanity-check, etc.) passed.

Engineering Impact

- Global / Societal: Enables reliable dual-sensor vision for search-and-rescue, firefighting, wildlife protection and critical-infrastructure inspection, even in resource-constrained regions.
- Economic: Removes need for costly mechanical calibration rigs; retrofit to existing payloads.
- Environmental: Fewer repeat flights → lower CO_2; supports SDG 13 – Climate Action.
- Ethical & Open Science: Fully open-source stack encourages peer review and responsible dual-use oversight.

Contemporary Issues

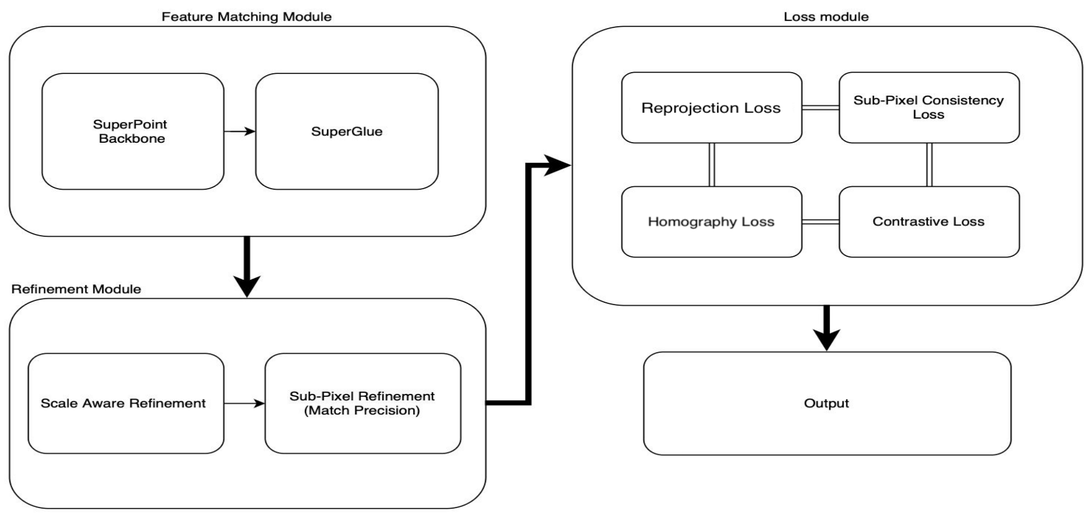
- Privacy: Drone imaging intersects GDPR / KVKK; on-device processing avoids raw data uplink.
- Dual-Use: Alignment tech has defense applications → export-control compliance required.

Key Technologies

Table 1: Core Technologies Utilized in *ThermalNet*

Category	Tool	Role
Deep CV Matcher	**SuperGlue**	Graph-NN correspondence learning
Differentiable Geometry	**Kornia**	Homography normalisation & losses
Optical Flow	**RAFT**	Sub-pixel keypoint refinement
Training Infra	**PyTorch + AMP**	40% speed-up

Figure 1: Architecture Diagram



CONCLUSION

- ThermalNet delivers real-time, pixel-level RGB–thermal alignment, validated by exhaustive tests, and is ready for pilot deployment on mobile platforms.

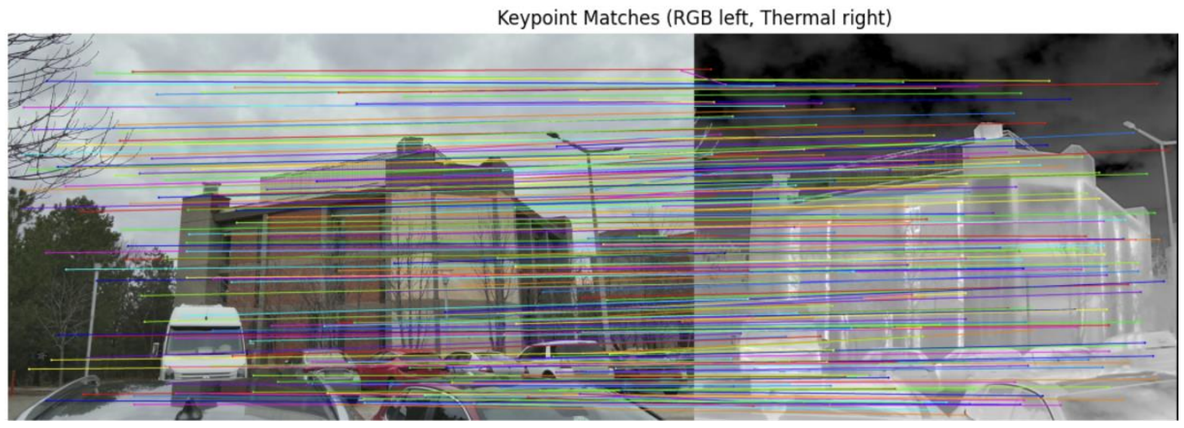


Figure 2: Sample matching

FUTURE WORK

- Curate additional urban & light-rain scenes to stress-test generalization.
- Add alignment-confidence score and inlier overlay toggle for field operators.
- Conduct live-flight trials with partner drone club to uncover edge cases.
- Environmental: Fewer repeat flights → lower CO_2; supports SDG 13 – Climate Action.
- Ethical & Open Science: Fully open-source stack encourages peer review and responsible dual-use oversight.

REFERENCES

1. Ö. Tuzcuoğlu, A. Köksal, B. Sofu, S. Kalkan, and A. A. Alatan, “XoFTR: Cross-modal Feature Matching Transformer,” Proc. CVPR Workshops (Image Matching Workshop), pp. —, 2024.
2. P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, “SuperGlue: Learning Feature Matching with Graph Neural Networks,” Proc. CVPR, pp. 4938–4947, 2020.

ACKNOWLEDGEMENTS

We wish to convey our deepest respect and heartfelt gratitude to Prof. Dr. Tolga Kurtuluş Çapın for his invaluable scholarly guidance, and to Mr. Fatih Demirtaş and the entire Raventech team for their generous support and insightful collaboration, all of which were indispensable to the success of this work.