Orbital Fingerprinting: Classifying LEO Debris Families with Proper Elements and Machine Learning

Michael Ling, Dr Yang Yang*
*Space Engineering Lecturer
School of Mechanical and Manufacturing Engineering
Email: yang.yang16@unsw.edu.au

UNSW SYDNEY

The problem: a crowded sky

The LEO environment is becoming increasingly dangerous.

- Over 36,500 debris objects >10 cm are tracked; millions are smaller.
- Fragmentation events (collisions, explosions) are the primary source of new, high-risk debris.
- **Challenge:** Identifying the origin of these fragments is critical for space situational awareness, attribution, and predicting future risks.

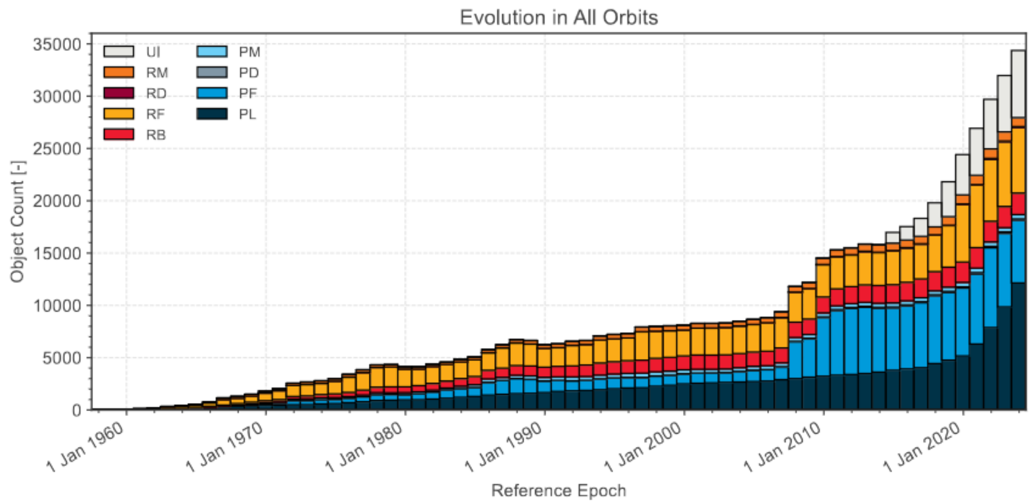

Figure 1: Evolution in RSO population since the beginning of human space activity (Credit: ESA Space Debris Office).

Prior work and research question

Limitations in the prior work:

- Classical methods: distance metrics + unsupervised clustering (e.g., DBSCAN) -> parameter sensitive, over-merge in dense LEO
- Existing ML based on fragment state difference features discard absolute orbital context, limiting discrimination
- Proper elements used singly; multi-set synergy underexplored

Question: How can supervised classification of RUOs (RSOs with unknown origin) improve using proper elements?

D. Wu, A. J. Rosengren, An investigation on space debris of unknown origin using proper elements and neural networks, Celestial Mechanics and Dynamical Astronomy 135 (2023) 44. doi:10.1007/s10569-023-10157-0.
A. Celletti, G. Pucacco, T. Vartolomei, Reconnecting groups of space debris to their parent body through proper elements, Scientific Reports 11 (12 2021). doi:10.1038/s41598-021-02010-x.

Research contributions

Contributions: Learned pairwise probabilistic similarity on multi proper element representations (modified equinoctial elements (MEE) + Poincaré (PNC) + Quaternion (QTN)) for robust debris family attribution

- End-to-end synthetic pipeline,
- Expanded feature sets across MEE + PNC + QTN (incl. QTN_p),
- Empirical ablations (1^{st}/2^{nd}-order terms; element-set combos).

Takeaway:

- Multi-representation proper elements (MEE + PNC + Quaternion) outperform hand-crafted orbital distance metrics for debris family attribution.
- Preprint and source code coming soon.

Pipeline overview

Parent ephemerides → stochastic breakup model (NASA EVOLVE 4.0) → high-fidelity orbit propagation → proper elements → neural network

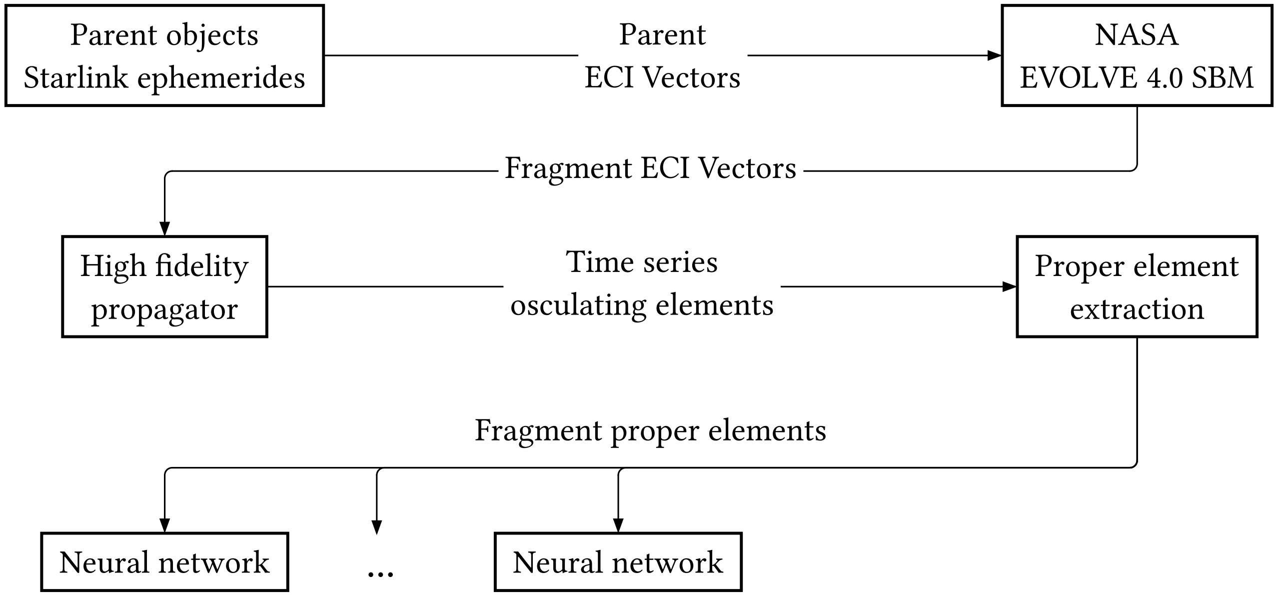

Figure 2 Compute pipeline.

Data & breakup model

- Starlink ephemerides + EVOLVE 4.0 stochastic breakup model (explosions).
- Sample → ΔV model → ECI states; filter immediate re-entries.
- Document fragment size & area-to-mass-ratio distributions.

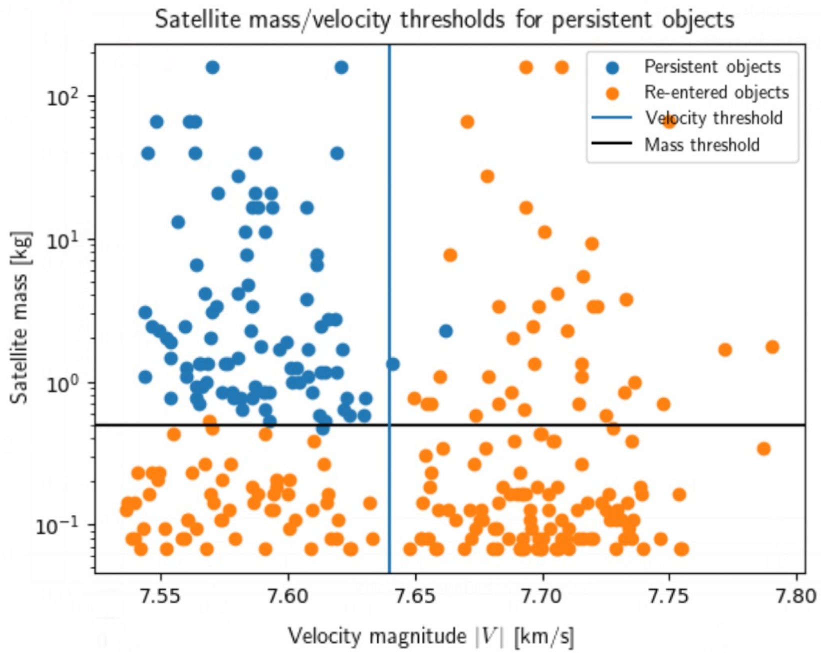

Figure 3 Filter immediate re-entries

Proper elements: what & how (1)

- Element sets: MEE (the following equation), PNC, QTN.
- Extraction via circle-fit; interval sampling for robustness.
- These are the debris-family "fingerprints".

$$h = e\cos(\bar{\omega}) \qquad k = e\sin(\bar{\omega})$$

$$p = i\cos(\Omega) \qquad q = i\sin(\Omega)$$

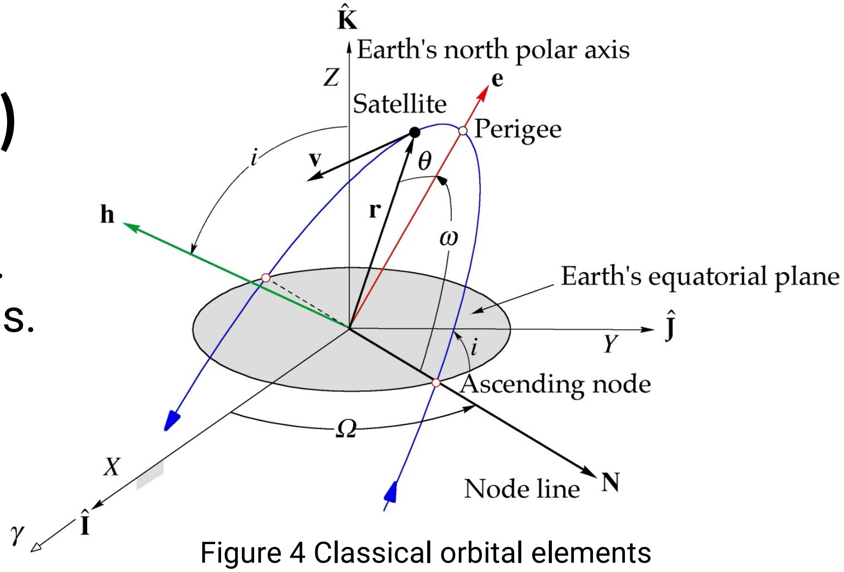

Figure 4 Classical orbital elements

TABLE III: Element sets selected contain pairs that allow formation of unit circle identities, suitable for proper element extraction

Element set	Pair 1	Pair 2	Pair 3 / Element
MEE	h, k	p, q	semi-major axis (SMA)
PNC	ξ, η	u, v	action-related SMA (Λ)
QTN	q_0, q_3	q_1, q_2	e_X, e_Y

Proper elements: what & how (2)

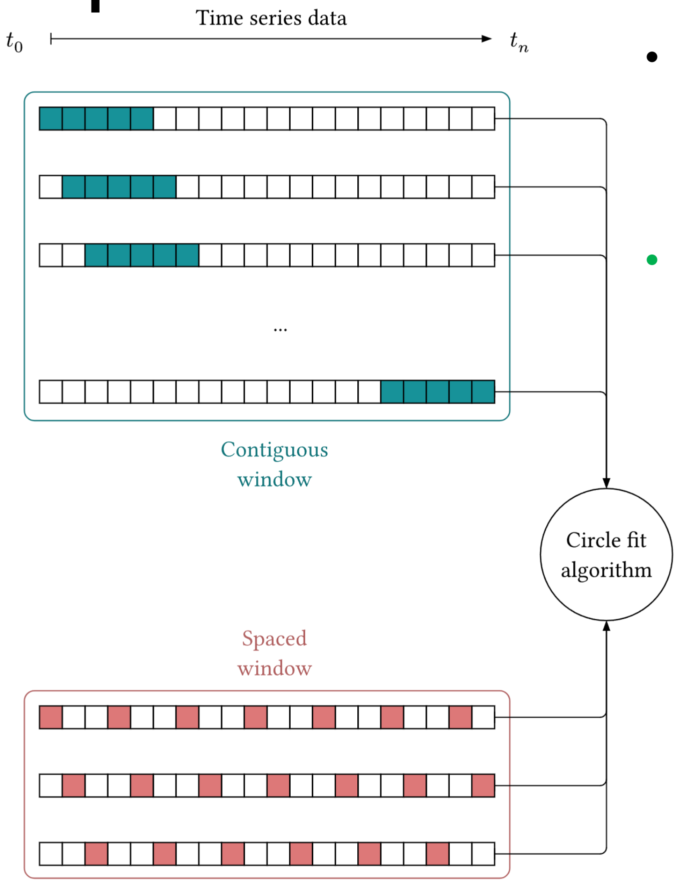

Figure 5 Contiguous window vs spaced window

- **Contiguous window sampling couples the fit too tightly to local, high-frequency perturbations,** making it sensitive to noise rather than revealing the true secular trend.
- Spaced window sample is introduced.

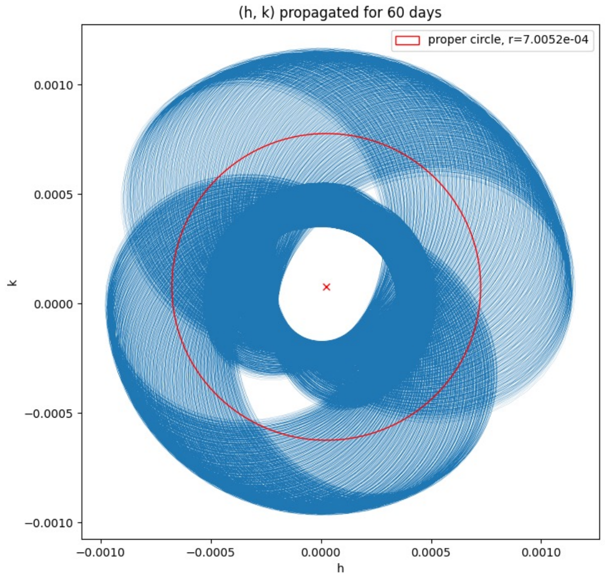

Figure 6 Circle fitting for MEE proper elements

Quality control on proper elements

- Bandwidth test + radial coverage bins reject bad fits.
- Cleaner inputs → better downstream classification.
- **Blue smears:** the raw time series of the modified equinoctial elements h, k.
- **Red circles:** the best-fit circles found by the extraction algorithm.
- **Green x:** the fitted circle centres (the “proper elements”).
- **Red shaded squares:** indicate that this fit *passed* the filtering test — i.e., was deemed a valid proper element extraction.

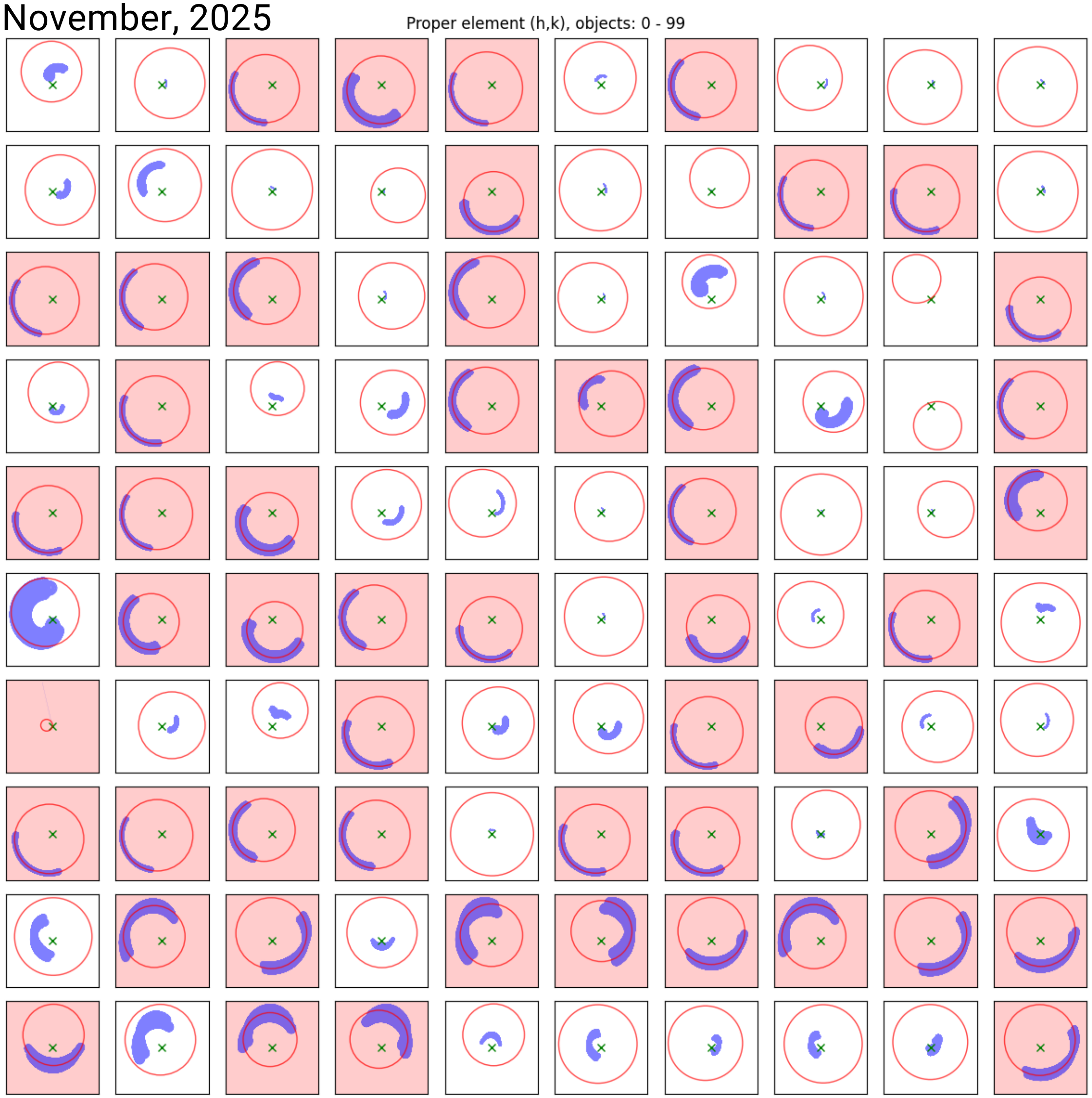

Figure 7: Results of filtering MEE proper elements h, k.

Dataset construction

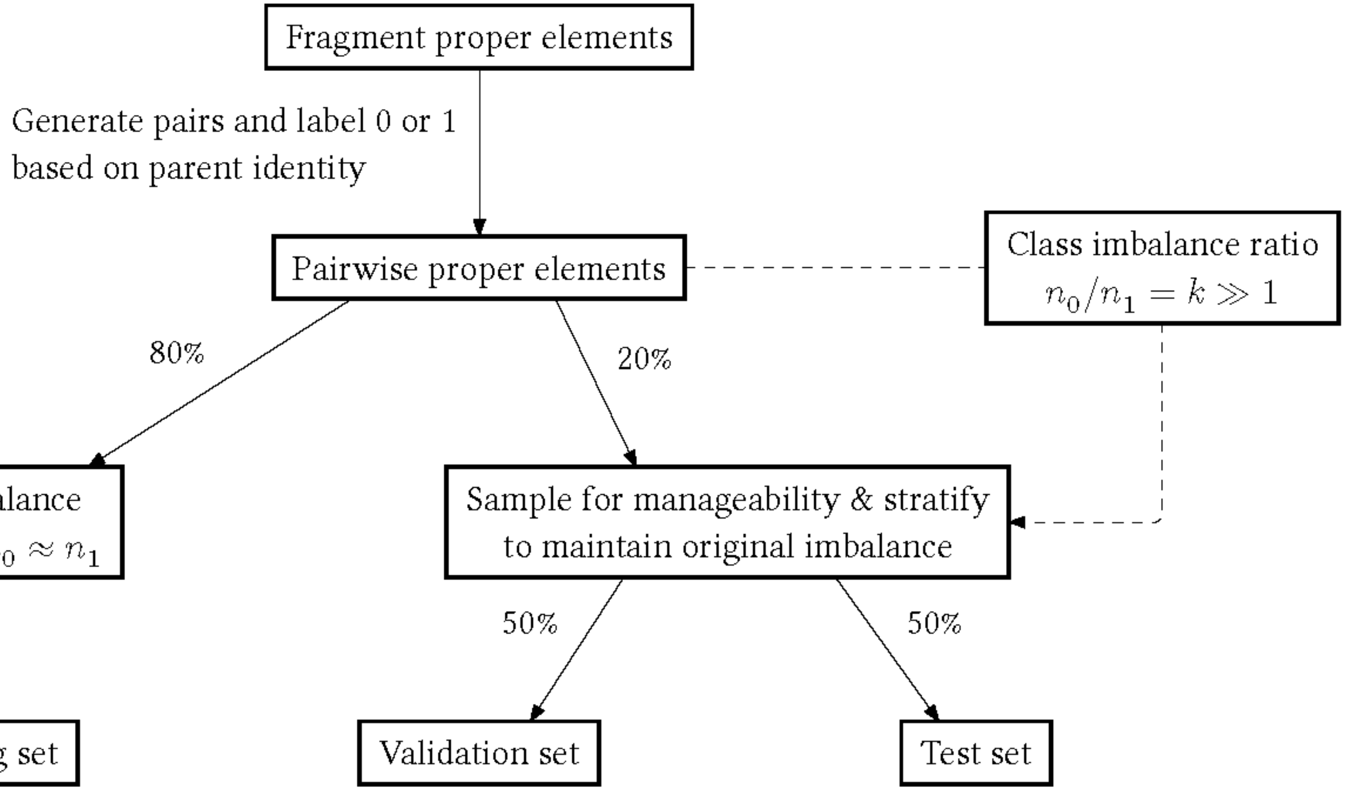

Figure 8 Dataset construction.

- Pairwise **proper elements** labelled: 1 = same parent, 0 = different.
- Strong class imbalance: $n_0/n_0 \gg 1$.
- **80 % → training** (rebalanced); **20 % → validation/test** (keep imbalance).
- **Preprocessing:** standardise features, sin/cos encode angles, normalise quaternions.

Neural network architecture

- **Input:** Pair of fragment proper elements.
- **Output:** P(same family |features) — probability of common origin.
- **Training:** Binary cross-entropy loss on labelled pairs (0/1).
- **Capability:** Learns nonlinear relationship in proper elements beyond fixed distance metrics.
- **Inference:** Apply probability threshold (e.g., F_1-optimal); analyse via ROC/PR trade-offs.

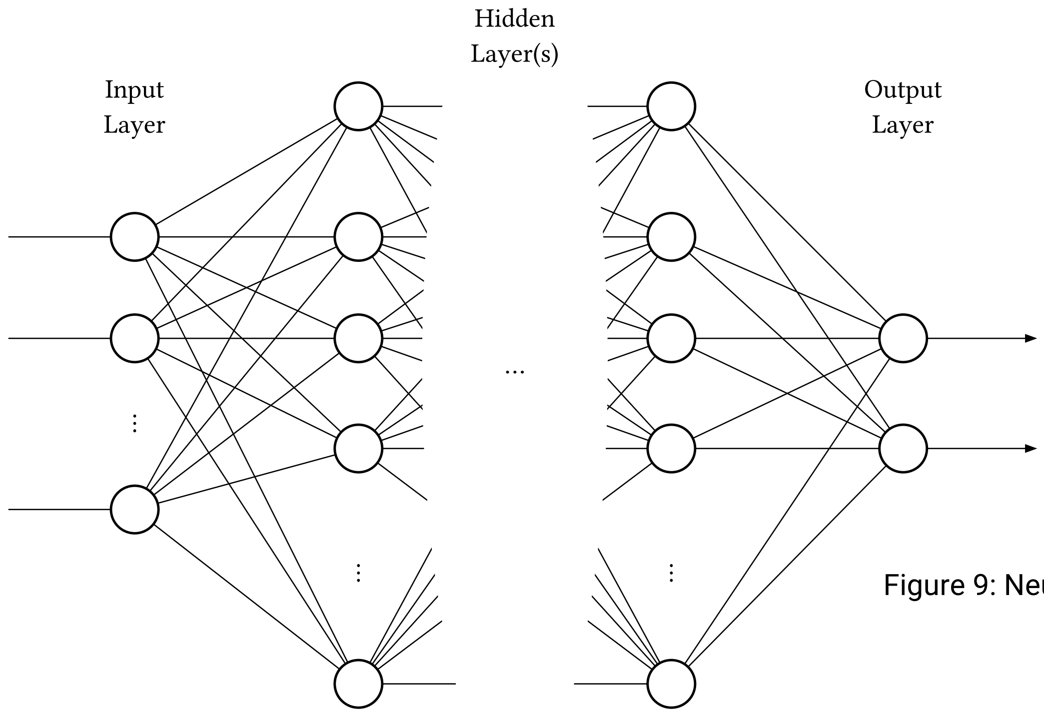

Figure 9: Neural network architecture.

Results I — confusion matrices (threshold = 0.5)

- Model output: For a fragment pair (i, j) the neural net returns $p = P(\text{same breakup family} \mid \text{features})$.
- Threshold = 0.5: baseline error balance in the confusion matrix; classify pair as same-family if $p \geq 0.5$.
- Adding element sets reduces errors; MEE+PNC+QTN best.

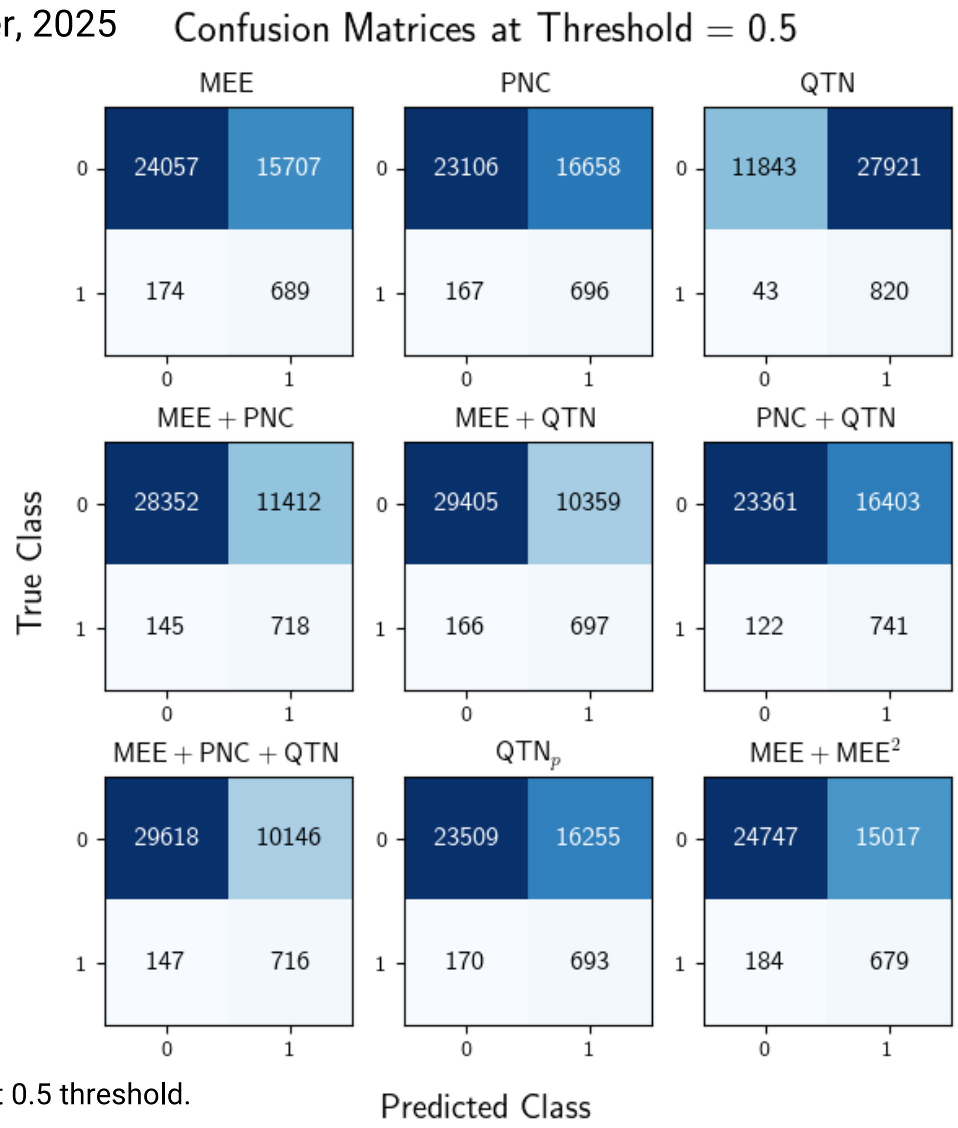

Figure 10: Confusion matrices for each model at default 0.5 threshold.

Results II — confusion matrices (optimal F_1 score)

- Model output: For a fragment pair (i, j) the neural net returns $p = P(\text{same breakup family} \mid \text{features})$.
- Optimal-F_1 threshold: Scan p thresholds (e.g., 0.05→0.95) on validation set; balancing precision and recall for the imbalanced dataset

$$\text{Precision} = \frac{TP}{TP + FP}, \text{Recall} = \frac{TP}{TP + FN}$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

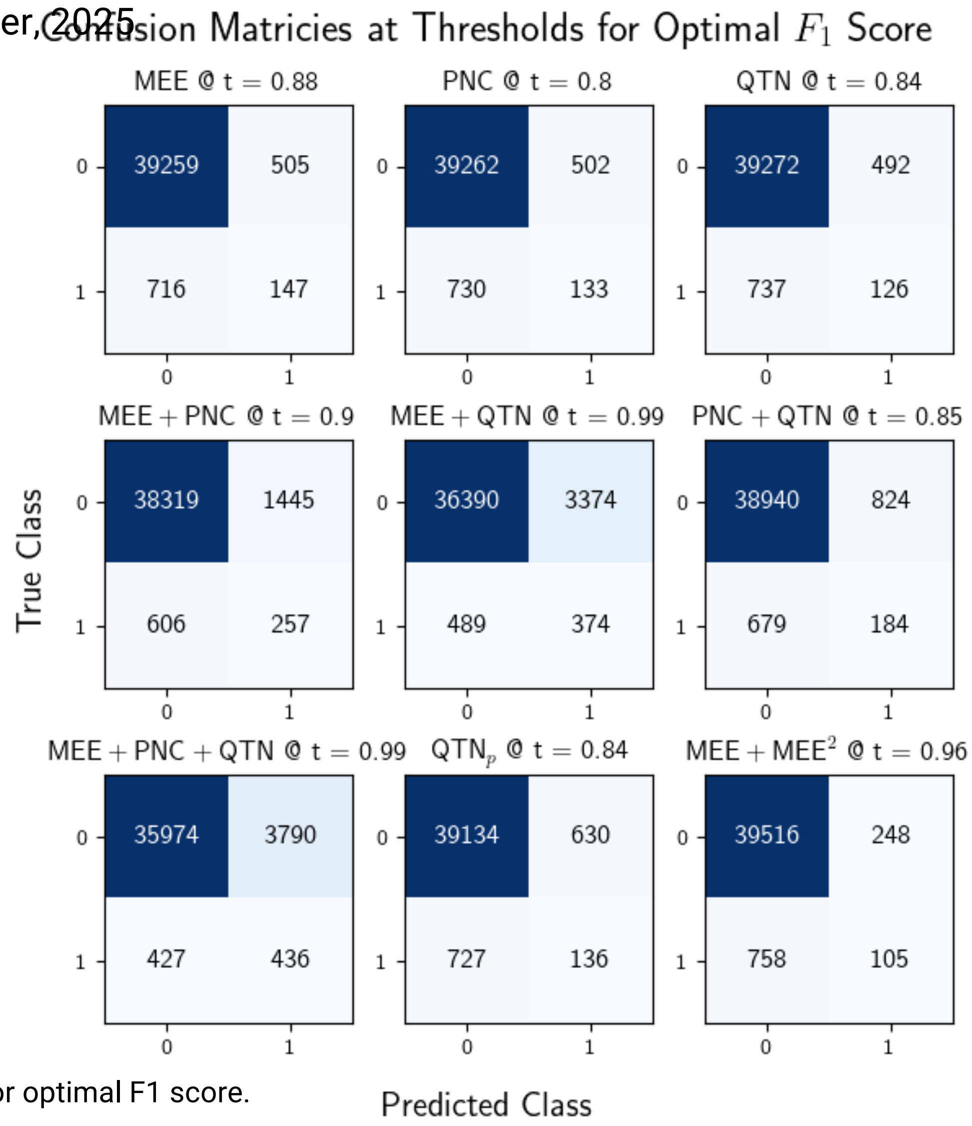

Figure 11: Confusion matrices for each model at thresholds selected for optimal F1 score.

Results III — receiver operator curves (ROC)

- The greater the area under the curve (AUC) the better performance.
- AUC: **triple-set leads (≈0.858)**.
- Weighted F_1/Accuracy highest for MEE+PNC+QTN; QTN_p > QTN; MEE^2 ≈ MEE.

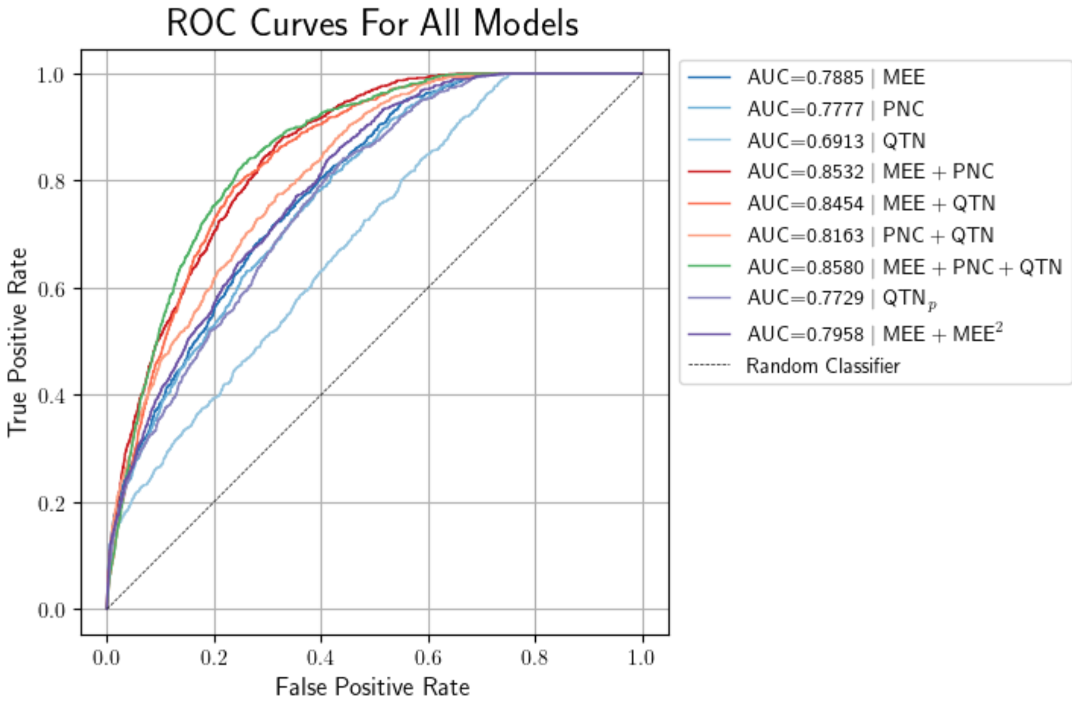

Figure 12 Overview of receiver-operator curves for all models.

Summary, takeaway & future work

- Orbital fingerprinting: Stable proper elements (MEE, Poincaré, QTN/QTN_p) enable robust debris family attribution.
- Pipeline: Modular end-to-end (synthetic breakup → propagation → proper element extraction → supervised classification).
- **Take-home**:
  - Trustworthy multiple proper elements → stronger family reconstruction; diminishing gains beyond two sets.
  - Origin classification on imbalanced data → ROC/PR-tuned thresholds improve recall-precision balance.
- Next step: Real fragmentation events, probabilistic breakup surrogates, etc.

Questions?

Michael Ling, Dr Yang Yang*
*Space Engineering Lecturer
School of Mechanical and Manufacturing Engineering
Contact: yang.yang16@unsw.edu.au

UNSW
SYDNEY

Backup - data & breakup model

- Starlink ephemerides + EVOLVE 4.0 stochastic breakup model (explosions).
- Sample $\rightarrow \Delta V$ model $\rightarrow$ ECI states; filter immediate re-entries.
- Document fragment size & AMR distributions.

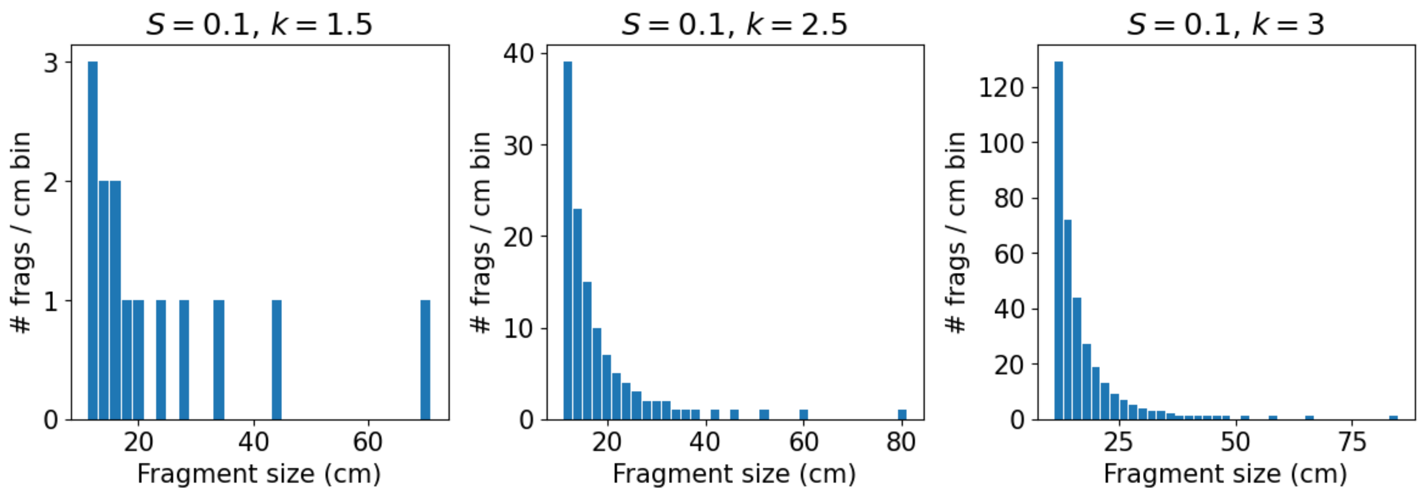

Figure 3 The granularity of fragmentation is shown here by varying the k parameter (k = 1.6 for explosions).

Backup - key insights & ablations

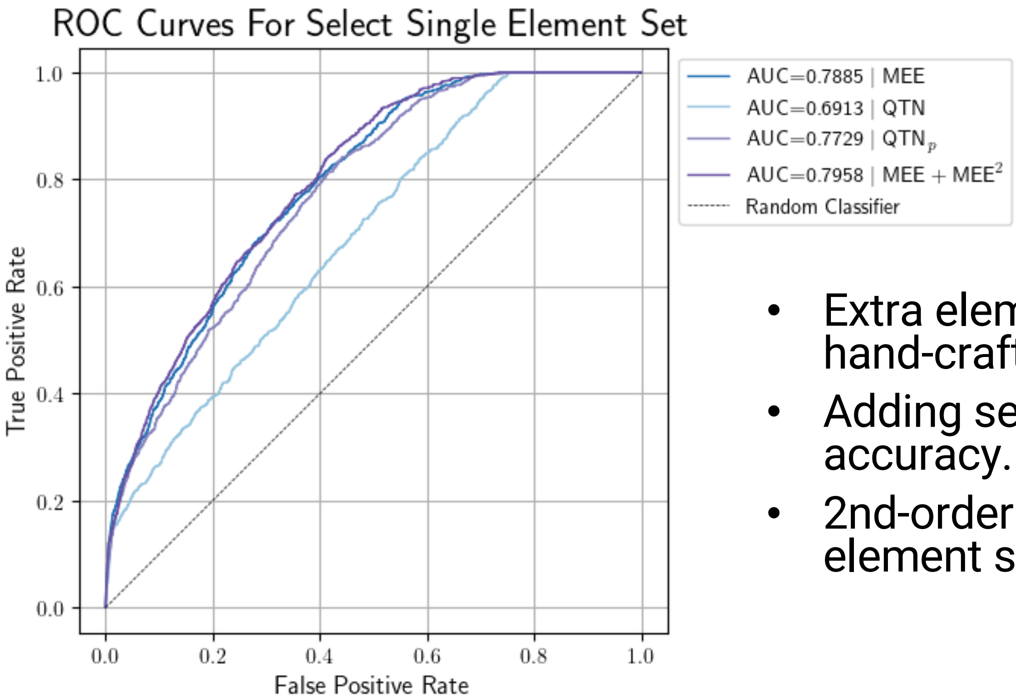

Figure 9: ROC-AUC plots for selected single element sets.

- Extra element sets boost performance more than extra hand-crafted tweaks.
- Adding semi-latus rectum p (QTN$_p$) restores QTN accuracy.
- 2nd-order terms yield small gains vs adding another element set.

Backup - limitations & validity

Synthetic data; class imbalance; threshold depends on use-case.

Environment drift remains; models must adapt to evolving distributions.