Machine Learning Classification of Micro Silica Glass Jewellery Beads for Thermoluminescence Dosimetry

Lukasz Tomaszewski [1,2], Caroline Shenton-Taylor [1]

l.tomaszewski@surrey.ac.uk , c.shenton-taylor@surrey.ac.uk

[1] Nuclear Applications and AI Group, Department of Physics, University of Surrey, Guildford, GU2 7XH, UK

[2] Nuclear Threat Reduction Network (NTR-Net), David Smith Building, 73 St. Michael's Hill, Bristol, BS2 8BH, UK

1. Background

- Thermoluminescence dosimetry (TLD) is a process of measuring ionising radiation exposure by analysing the emitted light from the heated irradiated material and is proportional to the absorbed radiation dose received (material specific), making it important for personal health monitoring within the nuclear and health sectors.
- Thermoluminescence (TL), or emitted light, is a result of trapped electron recombination, where trapped electrons can move between energy states within the material and released at certain temperatures, emitting light in the process; this is typically depicted by a glow curve of light intensity as a function of temperature.
- Machine learning (ML) can be applied to decipher hidden patterns in glow curves, exposing features: heating parameters and material properties. Allowing for absorbed dose calculations and corrections to be made.
- Due to the importance of TLD, in that it calculates personal radiation exposure, new dosemeter materials and analysis methods are critically sought after. Micro silica glass jewellery beads have shown to be a strong retrospective TL dosemeter producing stable and linear TL responses. The research undertaken explores the use of ML to expose the heating and material features of these glow curves, informing their use as a dosemeter.

2. Aims

- To explore if ML can accurately extract unique elements of micro silica glass bead glow curves (i.e colour, size) and imply hidden identifying markers.
- To establish the impact of gaussian-smoothing the raw glow curve spectra on classifying the unique glow curve parameters. Is there any loss?
- To inform the most suitable ML algorithm for classifying micro silica glass beads for TLD and develop enhanced assistive ML analysis techniques.
- To simulate labelling unspecified glow curves, in which material features can be revealed and absorbed dose calculations for nuclear forensics.

3. Method

4. ML Multi-Output Classification

- **Seed Dataset:**
  - 1,016 mass-normalised glow curve responses.
  - Sample types:
    - TLD-100 (LiF:Mg,Ti), TLD-100H (LiF:Mg,Cu,P) chips.
    - Toho & Mill Hill micro silica glass beads (Fig. 1).
  - 9 silica bead colours with 3 distinct physical sizes.
  - 2 input training datasets with raw and gaussian-smoothed ($\sigma = 2$) glow curves.
- **13 Features (per Glow Curve):**
  - **Temperature:** min, max, rate, pre-heat: temp, time.
  - **Sample:** colour, size, type.
  - **Irradiation:** exposed dose, time since, duration.
  - **Readout:** 2 Hoya U-340 filters: 2.5mm & 5mm.
- **ML Algorithms:**

  All models set to categorical multi-output architecture and independent feature model metrics.
  - Neural Network (NN)
  - Support Vector Machine (SVM)
  - Random Forest Classifier (RFC)
  - K-Nearest Neighbour (KNN)

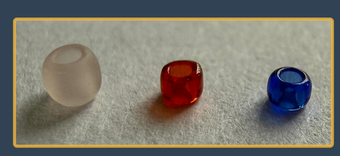

Figure 1: Three micro silica glass jewellery beads: 2 Toho (Japan): size 11 frosted rosaline (left), size 15 siam ruby (middle) and a Mill Hill (Czech Republic) size < Toho 15 blue.

4.1. Findings

- Sample size and colour are surprising higher than expected, indicating unique spectra elements.
- The Random Forest achieves the highest accuracy across all features and currently is the most suitable algorithm for this current seed dataset.
- Noise-sensitive ML algorithms (RFC & KNN) benefit most from the gaussian-smoothing preprocessing.
- The neural network underperforms, due to the limited dataset size and current dataset structure.
- All irradiation features show minimal predictive ability, indicative of lack of seed data variance.

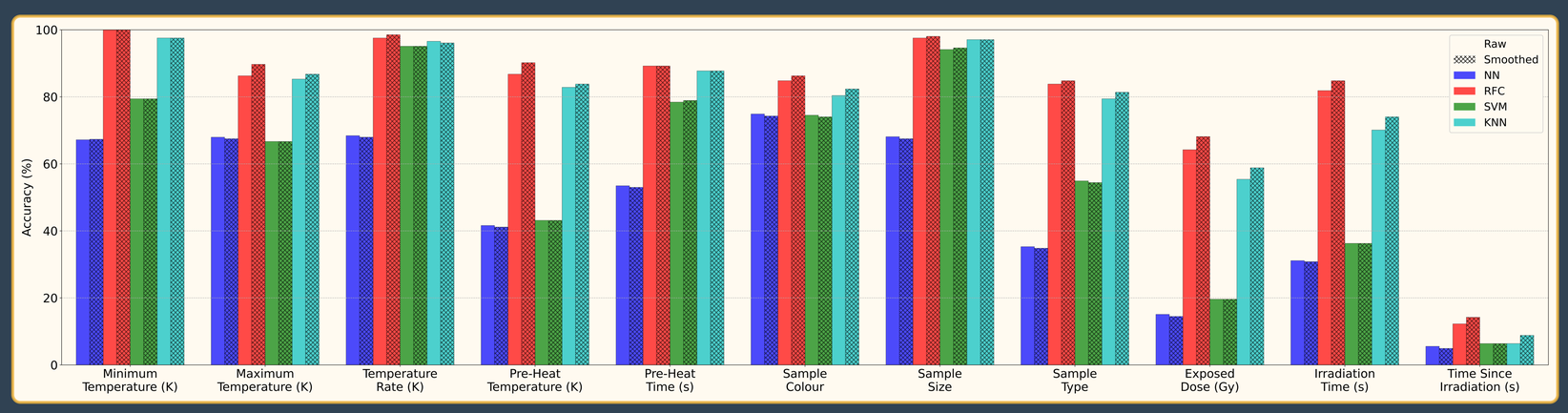

Figure 2: Accuracy scores per feature output of 4 ML algorithms (NN, RFC, SVM, KNN) using two training datasets consisting of 1,016 glow curve spectra with 13 unique features; the difference in the two datasets is one has the raw glow curve spectra and the other uses Gaussian-smoothed spectra. The results imply consistent and stable performance in the RFC and KNN ML algorithms, and feature accuracies comparable (in terms of physics) to the current research field, where "time since irradiation (s)" is a lost metric and harder to identify and good sample feature prediction: with a surprising result being strong classification of sample size (bead size). Additionally, the models mean metric scores for all algorithms in Table 1.

Mean Multi-Output Classification Results

ML Algorithm*	Mean Accuracy	Mean Loss**	Mean Precision	Mean F1	Mean Recall
NN (R)	0.52	0.48	0.36	0.41	0.52
NN (S)	0.53	0.47	0.35	0.41	0.52
SVM (R)	0.64	0.36	0.58	0.58	0.64
SVM (S)	0.64	0.36	0.58	0.58	0.64
RFC (R)	0.83	0.17	0.80	0.80	0.82
RFC (S)	0.84	0.16	0.83	0.83	0.84
KNN (R)	0.79	0.21	0.79	0.78	0.79
KNN (S)	0.81	0.19	0.81	0.80	0.81

*R: Raw Spectra, S: Smoothed Spectra **Hamming Loss

Table 1: Metric results of multiple ML algorithms, with raw or Gaussian-smoothed inputted spectra data that classifies 13 features unique to each glow curve. The tabulated data indicates the RFC and KNN showing the most stable results with negligible difference in input data.

5. Conclusion

- Due to the limited training dataset size, moderate model performance was expected and achieved.
- Neural network performance was poor, due to lack of data. Traditional data augmentation was explored but failed to achieve positive results due to unique feature destruction in the glow curve.
- The aims were achieved by classifying the sample characteristics (colour, size) and emphasising that ML can aid in TLD analysis for silica glass beads.
- The two inputted datasets showed negligible difference in mean results but favoured Gaussian-smoothing in individual feature analysis (Fig 2).

6. Future Work

- Explore unique data augmentation techniques for increasing training dataset and model accuracy.
- Implement ML generated synthetic glow curves to bolster the training dataset size to clear any dataset biases and improve data value variance.
- Bolster the seed dataset by extracting glow curves from more bead colours at varying doses.
- Extend "time since irradiation" feature data, to explore improving fading feature identification.
- Enhance the ML models by including isotope irradiation features, exploring isotope identification and radiation type (beta, gamma).

UNIVERSITY OF SURREY

Acknowledgements

This PhD research project is funded by the Nuclear Threat Reduction Network (NTR-NET).

Special thanks to S.Heisig, Dr I.Strudwick and J-W.Brown of the University of Surrey's Department of Physics, radiation laboratory technical team.

© British Crown Copyright 2026 / AWE