REIMS in Beef

Rapid detection and specific identification of offals within minced beef samples utilising ambient mass spectrometry [1]

Jesse Wood[1]

[1]School of Engineering and Computer Science — Te Kura Mātai Pūkaha, Pūrorohiko
Victoria University of Wellington — Te Herenga Waka

Literature Review, October 2022

Table of Contents

Background

- Criminals add stuff to meat products (adulteration) for economic gains.
- Meat adulteration in non-meat products of $<1\%$ expected (and allowed) as it is considered cross-contamination, and not for economic gains.
- Adulterations levels from (15%-20%) are considered criminal as they are likely for economic gains.
- 2013 European Horsemeat scandal is an example of this.
- In response, European Union (EU) declared that non-meat offal cuts must be declared on product labels.
- Recent study [2] in the UK ($n=665$), found $>1/5$ of samples contained non-declared meat species.
- E.g., for the 2013 European horsemeat scandal, REIMS could detect the adulteration, and identify that adulterant as horse.
- Rapid evaporative ionization mass spectrometry (REIMS)
- Minced beef products are often ready-to-go, and pre-cooked, so a method is needed that works on raw/cooked meat products.

Motivations

- DNA sequencing can only differentiate between different species, not offal adulteration from the same species.
- Our work [3] showed Gas Chromatography can differentiate between species easier than the individual parts of a fish.
- Vibrational spectroscopy can detect adulteration, but not the specific offal present.
- Both DNA methodologies and vibrational spectroscopy are ineffective at detecting these adulterations.
- Traditional chromatography/mass spectrometry hasn't been tried, due to the time to prepare/analyze samples.
- Ambient Mass Spectrometry (AMS) has the potential to identify unique/significant metabolites. GC-MS cannot do this!
- Significant Markers (or important variables) are ions that are unique to a specific offal cut, and present in all samples.
- Looking for a reliable, accurate and rapid method that can detect criminal adulteration levels in beef.

Data I

- Cheap offal products can be added to beef tissues when they are minced in food processing to cut corners and increase profits.
- Minced beef (1 class) with adulteration from beef brain, heart, kidney, large intestine and liver tissues (5 classes).
- Outliers are hybrid spectra - a homogenous mix of beef and adulteration - at a given adulteration level (i.e. 20%, 10%, 5%, 1%).
- Method facilitates real-time classification, with classification output produced every second.
- METLIN metabolites database, and LIPID MAPS can provide annotated labels for spectra.

Remark

Existing databases of lipid spectra are a source of domain knowledge that can be transferred to new tasks to improve performance and create semantically meaningful features.

Data II: Preprocessing + Postprocessing

Pre-processing (before PCA-LDA):

- Prototype abstract model builder
- Masslynx pre-processing algorithms
- Background subtracted
- Lockmass corrected
- Normalized by TIC (total ion count)

Post-processing (after PCA-LDA):

- Mean-centered
- Pareto scaled
- Grouped by class

Method I

In [1] they propose REIMS for detecting beef adulteration.
Metrics:

- R^2 measures the variation in samples.
- Q^2 measures the classification accuracy.
- *RMSE-CV* measure cross-validated root means squared error.

Feature Selection:

- Variable Importance Projection (VIP)
- S-plots?

Remark

Applies trivial statistical modelling to state-of-the-art chemistry techniques, i.e. no state-of-the-art AIML.

Method II

- Chemometric analysis (VIP + S-plots) of REIMS could detect unique/significant markers.
- Principal component analysis linear discriminant analysis (PCA-LDA) [4] using orthogonal partial least squares discriminant analysis (OPLS-DA) [5].
- PCA-LCA used for dimensionality reduction - classification, respectively.
- Detect outliers based on standard deviation outside 20σ of the mean for any class.
- They provide a very detailed description of their method from the chemistry side, including instruments and their settings. Good for reproducibility and understanding.

Remark

Their PCA-LDA model had many hyperparameters (e.g. # principal components, 20σ threshold for outliers, mass range (m/z) for REIMS) by the authors via trial and error. These are nuisance variables.

Results

- PCA/LDA (with manual hyper-parameter tuning) can effectively detect adulteration - i.e. cluster different classes within adulteration levels (i.e. 15-20%).
- The adulteration levels were measured on raw/boiled minced beef.
- Raw: brain (5%), heart (1-10%), kidney (1-5%), large intestine (1-10%), liver (5-10%).
- Beef and large intestine were too similar to detect outliers with PCA-LDA. Perhaps very similar tissue composition.
- Within adulteration levels (i.e. 15-20%), their model can predict adulteration with perfect precision, i.e., all predicted adulterations were correct.
- Boiled: brain (5-10%), heart (1-10%), kidney (1-5%), large intestine (1-10%), live (5-10%).
- Boiled samples are harder to classify. More principle components were needed to correctly identify adulteration for boiled samples.

Why it matters?

- REIMS is a cheap and rapid method for detecting adulteration in minced beef in a factory setting.
- REIMS can detect both adulterations, and the specific adulteration present, superior to other methods.
- Many meat products are pre-cooked, REIMS detects adulteration (at criminal levels) in raw/boiled meat.
- REIMS can provide a paradigm shift across many authenticity applications.
- Previous work [6] (from the same author) shows can be successfully applied to fish REIMS data.

Limitations

- Manual hyperparameter tuning (e.g. # principal components, threshold for outliers, mass range) can be automatically selected, or replaced by models that don't need them at all!
- Basic dimensionality reduction techniques (e.g. PCA) were used. Future work should consider t-SNE [7].
- Basic supervised statistical models (e.g. LDA, OPLS-DA) was used for classification. Future work should consider CNNs [8, 9], GANs [10], Diffusion [11].
- Potential for transfer learning (incorporate previously existing data) to improve performance for few-shot classification tasks.

Related Works

- [6] use REIMS for fish fraud detection.
- [2] Recent study in the UK (n-665), found $>1/5$ of samples contained non-declared meat species.
- [12] shows cross-species adulterations (horse-meat) can be detected at levels greater than 1%
- [10] shows that Generational Adversarial Networks (GANs) can be trained to detect anomalies when the reconstruction error exceeds a threshold, very similar to the statistical analysis provided here [1].

[1] C. Black, O. P. Chevallier, K. M. Cooper, S. A. Haughey, J. Balog, Z. Takats, C. T. Elliott, and C. Cavin, "Rapid detection and specific identification of offals within minced beef samples utilising ambient mass spectrometry," *Scientific reports*, vol. 9, no. 1, pp. 1–9, 2019.

[2] H. Mackay, "Meat testing: A fifth of samples reveal unspecified animals' dna," Sep 2018. [Online]. Available: https://www.bbc.com/news/uk-45371852

[3] e. a. Wood J, Nguyen B, "Automated fish classification using unprocessed fatty acid chromatographic data: A machine learning approach," *AI 2021: Advances in Artificial Intelligence*, vol. 35, no. 52, pp. 12 536–12 544, 2022.

[4] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[5] J. Boccard and D. N. Rutledge, "A consensus orthogonal partial least squares discriminant analysis (opls-da) strategy for multiblock omics data fusion," *Analytica chimica acta*, vol. 769, pp. 30–39, 2013.

[6] C. Black, O. P. Chevallier, S. A. Haughey, J. Balog, S. Stead, S. D. Pringle, M. V. Riina, F. Martucci, P. L. Acutis, M. Morris *et al.*, "A real time metabolomic profiling approach to detecting fish fraud using rapid evaporative ionisation mass spectrometry," *Metabolomics*, vol. 13, no. 12, pp. 1–13, 2017.

[7] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.

[8] K. Bi, D. Zhang, T. Qiu, and Y. Huang, "Gc-ms fingerprints profiling using machine learning models for food flavor prediction," *Processes*, vol. 8, no. 1, p. 23, 2019.

[9] D. D. Matyushin and A. K. Buryak, "Gas chromatographic retention index prediction using multimodal machine learning," *Ieee Access*, vol. 8, pp. 223 140–223 155, 2020.

[10] F. Di Mattia, P. Galeone, M. De Simoni, and E. Ghelfi, "A survey on gans for anomaly detection," *arXiv preprint arXiv:1906.11632*, 2019.

[11] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," *arXiv preprint arXiv:2010.02502*, 2020.

[12] Y.-H. P. Hsieh and J. A. Ofori, "Detection of horse meat contamination in raw and heat-processed meat products," *Journal of agricultural and food chemistry*, vol. 62, no. 52, pp. 12 536–12 544, 2014.