I. Introduction

The recycling of scrap in Europe remains a secure, sustainable, and cost-effective source of raw materials despite political conflicts with mining countries. This approach helps prevent conflicts between the local population and the mining industry. Online classification of recycling materials is crucial for aligning them with standards. In digitized recycling, precise online testing of heterogeneous materials allows optimal control over the composition of recycled products, enabling alloy-to-alloy recycling. This targeted strategy promotes efficient use of scrap metal, contributing to a more sustainable metal production. However, non-destructive metrological solutions for copper and aluminium production are currently lacking. Assessing scrap relies on surface-based techniques, but these are limited to homogeneous materials or involve impractical sample collection and preparation, hindering online applications.

PGNAA measuring systems offer integral, non-destructive elemental analysis of complex material flows or batches. By recording neutron-induced gamma spectra, the method provides insights into the complete elemental composition of the material. The demonstrator facility used a neutron generator to produce monoenergetic neutrons, enabling thermal interactions for elemental analysis. PGNAA proves robust in analyzing large-volume samples and material flows, addressing challenges like inhomogeneity, coatings, or impurities. Its versatility extends to applications such as oil exploration, explosives detection, and online quality assurance in the cement and coal industries.

In the domain of spectroscopy, Machine Learning (ML) is integral to various techniques, including FTIR, NIR, MIR, Raman, ARPES, LIBS, and PGNAA, with applications ranging from distinguishing materials like tablets, juices, and wines to exploring iron ore deposits, analyzing atmospheres of extrasolar planets, detecting radioactive materials, and ensuring nuclear safety. The integration of ML into spectroscopic analyses enhances capabilities across scientific and industrial domains.

Building on our prior research in non-destructive spectral classification using PGNAA, with a specific focus on copper and aluminum alloys, we have achieved promising results. Our approach employed maximum log-likelihood and CNN (Convolutional Neural Network) techniques. Additionally, for data generation, we utilized a probability distribution-based sampling method derived from the prolonged measurement of a spectrum [1, 2].

The classification of *mixed* copper alloy compositions with their increased similarity represents an even greater challenge in comparison to previously explored classification of non-blended alloys. This investigation is the first contribution of this paper. The investigation of metal alloy proportions holds particular relevance in recycling processes. Utilizing gamma spectra, we emphasize online classification and employ PGNAA for non-destructive material analysis. For the first time, we apply Neural Networks (NN) in the classification of metal alloys, alongside CNN and maximum log-likelihood. Our findings demonstrate a significant advancement, surpassing the current state-of-the-art in CNN classification of copper alloys and achieving a superior classification rate in just one-fifth of the time.

II. Data & PGNAA

PGNAA continuously analyzes materials through neutron beam irradiation, exciting atomic nuclei and inducing various nuclear reactions. The excited nuclei relax by emitting gamma quanta, producing a characteristic gamma radiation in the sample material. This radiation is measured with a gamma ray spectrometer, and the gamma energies are visualized exemplified by Fig. 1. All PGNAA spectra used in this paper were acquired with an HPGe detector[1].

Fig. 1 depicts two spectra labeled as '2' and '10'. The detailed composition can be observed in Table I. The classification of metal alloys becomes more challenging with shorter measurement times due to increased noise and reduced individual energy measurements, as illustrated in Fig. 2.

Table I: The labeling of mixed alloys (from 1 to 10).

Labels	Materials	Labels	Materials
1	66% Cu1 + 33% Cu3	6	50% Cu1 + 50% Cu3
2	33% Cu2 + 66% Cu3	7	50% Cu2 + 50% Cu3
3	66% Cu2 + 33% Cu3	8	33% Cu1 + 66% Cu2
4	33% Cu1+33% Cu2+33% Cu3	9	66% Cu1 + 33% Cu2
5	50% Cu1 + 50% Cu2	10	33% Cu1 + 66% Cu3

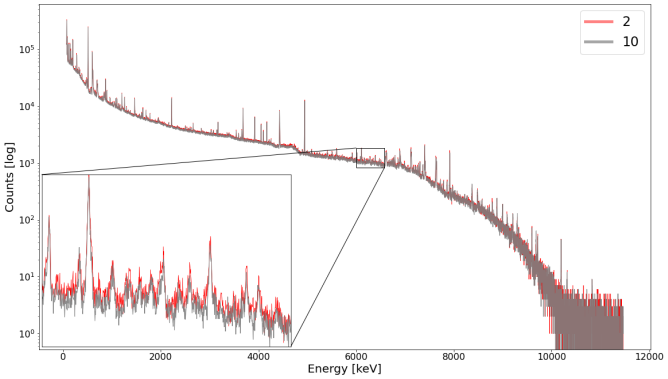

Fig. 1: Two copper alloy mix spectra, each measured for 2 h.

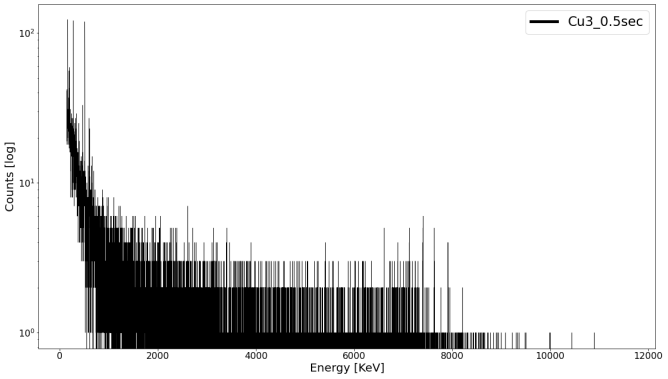

Fig. 2: Simulated Cu3 alloy spectrum, 0.5 sec measurement.

[1]You can access many metal alloys and other materials online: https://www.kaggle.com/datasets/smartfactoryowl/metalclass

III. Data Creation

The sampling method utilizes a fully measured spectrum as a representative dataset, treating it as a probability distribution through normalization. The discrete distribution derived from this spectrum captures new data (see [2]). During a long-term measurement (2-4 hours), approximately 70 million energy values are recorded. The probability, obtained by dividing the count rate of a specific energy value by the total count rate sum, converges according to the laws of probability theory (law of large numbers) towards the true distribution of count rates from the spectrum. This probability guides the rationale for sampling short-term measurements. In our scenario, 1 second is equivalent to 50,000 counts for all materials, and the sum of absolute frequencies is proportional to the desired measurement time.

IV. Results

For NN and CNN classification, 10 alloy mixtures (listed in Table I) were utilized, with 40000 training samples and 10000 testing samples per material, each comprising 100000 counts (2 seconds).

A. Convolution Neural Network

With CNN, we achieved a classification accuracy of 87.95%. We thus obtain the same result as in the current state of research with a measurement time that is 5 times shorter (see [1]). This is accomplished by converting the one-dimensional count rate of spectra into a matrix. We split the spectra at every 511th position, chosen as the escape peak, into 32 segments and include them in a matrix, resulting in a 32x511 matrix.

The CNN has the following architecture: It consists of four convolutional layers, sequentially increasing the number of channels from 1 to 3, 9, 12, and finally 9, each employing a 3x3 kernel, followed by batch normalization, rectified linear unit (ReLU) activation, and max-pooling with 1 padding. Subsequently, a single convolutional layer reduces the channels from 8 to 5 using a 2x2 kernel with ReLU activation. The network concludes with one fully connected layer. The output layer comprises 10 neurons. The total number of trainable parameters is 2476482. The model was trained using the CrossEntropyLoss, a learning rate of 0.001, and over 150 epochs, with the Adam optimizer.

B. Neural Network

The following NN was developed for classification and we achieved a classification accuracy of 92.57 %.

The input dimension has 16384 features and an output of 10 neurons. The network comprises 13 hidden layers of varying sizes, starting at 7000 neurons and decreasing to 10 neurons. In total, there are 264383165 trainable parameters. The model was trained using CrossEntropyLoss, a learning rate of 0.01, and over 150 epochs with the Adam optimizer.

C. Maximum log-likelihood

For the classification of short-time spectra 's' using the maximum log-likelihood method, we first obtain probability distributions for long-time spectra 'S_i', as described in Chapter III. The long-term spectra $(S_1, ..., S_{10})$ are now transformed into probability distributions. The classification process involves determining which distribution best fits the short-time spectra 's'. This task is efficiently handled by the maximum log-likelihood method. The method evaluates the likelihood of a short-time measurement given a completely measured spectrum, assigning it to the distribution that maximizes this likelihood. Evaluation of the short time measurement with this method with corresponding probability distribution yields $\max_i \log(p(s|S_i))$ (see [2]).

With the maximum log-likelihood, we achieve an impressive accuracy of 95.76% within a quarter of the counts, i.e., 25,000 counts (0.5 seconds). Maximum log-likelihood achieves better accuracy than NN and CNN. Furthermore, NN achieves better accuracy than CNN after approximately 0.25 seconds, as detailed in Fig. 3.

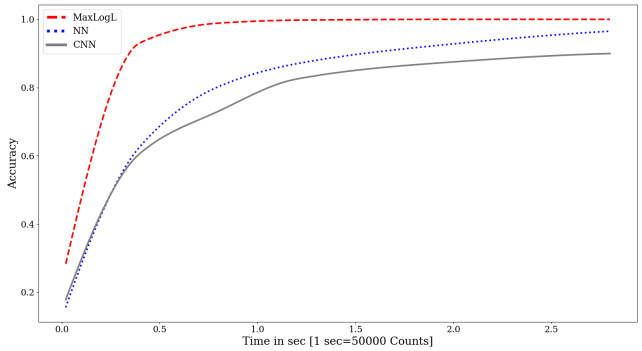

Fig. 3: Comparison of accuracy over time for Maximum Log-Likelihood (MaxLogL), NN and CNN methods.

V. Conclusion

In conclusion, our study makes a significant contribution to the field of non-destructive online classification of copper alloy mixtures using PGNAA. By employing maximum log-likelihood, we successfully classify gamma spectra nearly perfectly in less than one second, providing insights into the proportions of metal alloys-an aspect not explored before. Notably, we achieve improved CNN classification of copper alloys with a better classification rate in just one-fifth of the time compared to existing benchmarks. Additionally, our study demonstrates that NN achieve better accuracy than CNN. Our results highlight the possibility of distinguishing between mixed materials, even when they consist of alloys with similarities.

VI. Acknowledgments

The MetalClass project is funded by the German Federal Ministry of Education and Research with the funding reference number 01IS20082A/B. The copper alloys were provided by Wieland-Werke AG. The authors would like to thank Dr. Kai Krycki and Dr. Stefan Theobald for support and discussions.

References

[1] K. Y. Cheng "Prompt-Gamma Neutron Activation Analysis (PGNAA)' Metal Spectral Classification using Deep Learning Method' in *23rd Virtual IEEE Real Time Conference* (1.-5. Aug. 2022), 2022 pp. 6.

[2] H. Shayan 'PGNAA Spectral Classification of Metal with Density Estimations' *IEEE Transactions on Nuclear Science* (Volume: 70, Issue: 6, June 2023).

66% Cu1 + 33% Cu3	50% Cu1 + 50% Cu3	66% Cu2 + 33% Cu3	33% Cu1+33% Cu2+33% Cu3	50% Cu1 + 50% Cu2
33% Cu2 + 66% Cu3	50% Cu2 + 50% Cu3	33% Cu1 + 66% Cu2	66% Cu1 + 33% Cu2	33% Cu1 + 66% Cu3