On Feature, Classifier and Detector Fusers for ^{235}U Signatures Using Gamma Spectral Counts

Nageswara S. V. Rao*, David Hooper*, Jennifer Ladd-Lively
*Oak Ridge National Laboratory, Oak Ridge, TN, USA, {raons,hooperda,laddjl}@ornl.gov

***Abstract*—Three types of information fusion strategies are studied to assess the performance of classifiers for detecting low-level ^{235}U radiation sources, using features obtained from gamma spectra of NaI detectors. These strategies are based on using two spectral region features, fusing eight classifiers of diverse designs, and fusing multiple detectors located at different positions around the source. The inner, middle and outer groups of detectors, within a formation of two concentric circles and a spiral of 21 detectors, are identified based on their distance to the source, which is located at the center. This study provides two main qualitative insights into this classification task. First, the fusion of detectors leads to an overall improved classification performance, least in the inner group, most in the outer group, and in between for the middle group. Second, several classifiers and fusers achieve lower training error which does not translate to lower generalization error, indicating their over-fitting to training data.**

***Index Terms*—classifier fuser, detector fuser, radiation signatures, multiple detectors, gamma spectrum.**

I. Introduction

Classifiers for detecting low level radiation sources using spectral measurements from portable gamma-ray detectors have been studied under various scenarios. They include general formulations [15], detector networks [2], [16], use of NaI detectors [14], and deployments of multiple detectors [18] (to name a few). The signatures of low-level ^{235}U sources are of particular importance in nuclear safeguards and security tasks [7], [8]. The hand held gamma spectral detectors, such as NaI detector by Passport Systems and CsI detector by Kromek, are used in the field, and a variety of solutions have been developed to use their spectra to infer the presence of radiation sources. The underlying detection task is challenging because the Poisson distributed gamma measurements of these sources are hard to distinguish from the background measurements.

Our goal is to study the performance of Machine Learning (ML) classifiers combined with information fusion strategies involving features, classifiers and detectors, for this task. Our work builds upon recent ML studies using single detectors with a single ^{235}U spectral signature in [12], which revealed large variations in identically produced detectors and over-fitting by non-smooth ML methods. We expand this study by considering the information fusion methods involving two features based on gamma spectral regions, detectors located at different positions around the source, and the classifiers designed using diverse approaches, namely smooth and non-smooth, statistical and structural, and hyper parameter methods. In general, classification problems related to ^{235}U and other sources have been studied using data-driven ML methods [12], [3], [5], which complement the statistical and first principle methods [8], [7].

We utilize data sets collected under controlled conditions using 21 NaI detectors deployed over a 6 x 6 meters area in a formation of two concentric circles and one spiral, with the source located at the center. The activity levels in two spectral regions associated with ^{235}U signatures are estimated as counts at 1 second intervals, which are used as classifier features. These detectors form the inner, middle and outer groups based on their distance to the source, which represent an increasing degree of difficulty for classification. Eight different classifiers and six classifier-fusers are tested using measurements collected over multiple experimental runs.

A *detector profile* of a classifier is its classification error expressed as a function of the detector distance from the source. A *classifier profile* of a detector(s) specifies the errors of various classifiers and their fusers that use its(their) measurements. We estimate these profiles under the three fusion strategies using multiple data sets. They indicate an overall trend of decreasing error as detectors closer to the source are used and more detectors are fused – both phenomena are explained by the increased "effective" NaI capture area for gamma radiation. They also reveal unexpected trends of lower training error not being reflected in lower testing error for certain classifiers and fusers; this is more pronounced with two features but less with larger training sets. This study provides a concrete case of over-fitting by complex classifiers, whose training performance could be a misleading indicator of their generalization. Interestingly, unlike the other case, there does not seem to be a physics-based explanation to expect their superior performance, but one might be needed to guide their use in applications.

The organization of this paper is as follows. The datasets are briefly described in Section II. Experimental and analytical results of classifiers and fusers are described in Sections III and IV, respectively. The performance of combining features from multiple detectors is described in Section V. A summary and directions for future work are described in Section VI.

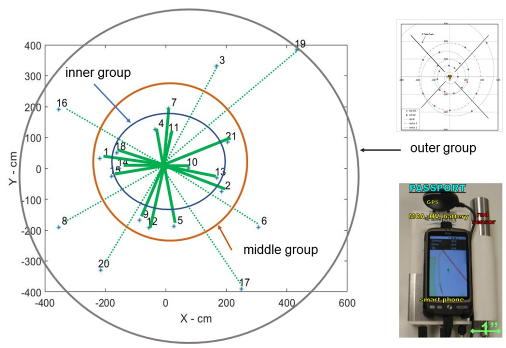


Fig. 1. Detector configuration of two circles and a spiral.

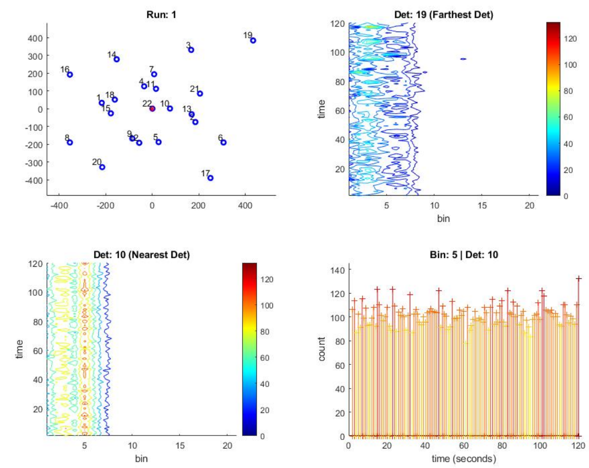


Fig. 2. Measurements dashboard: top left: detector locations; bottom left: spectrum of detector 10 closest to source; top right: spectrum of detector 19, farthest from source; and bottom right: counts per second at detector 10.

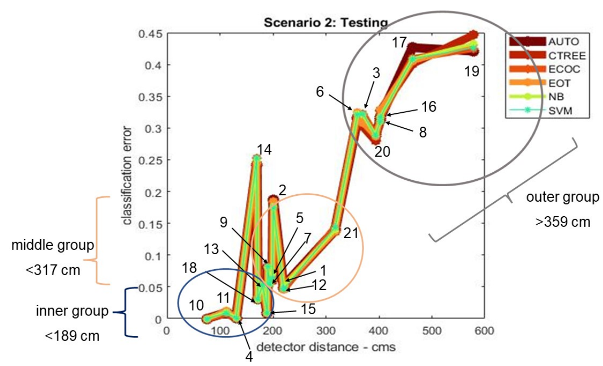


Fig. 3. Select-six classifiers have similar detector profiles.

II. Measurements

Measurements from controlled tests conducted at the Savannah River National Laboratory using ^{235}U source of 191 uCi strength are used in our study [13]. The source is introduced at the center of the Low Scatter Irradiator facility via a shielded conduit, and is stationary during the tests, and the detectors are arranged in the fixed geometric pattern shown in Fig. 1, which consisted of two rings with radii 2 and 4 meters with 4 and 5 detectors, respectively, and a spiral of 12 detectors. Each detector employed a 2" x 2" x 11" NaI sensor connected to a mobile phone, and an example is shown in Fig. 1. All detector are identically produced and configured, and the quality of their measurements is a reflection of their NaI sensor material. The nearest and farthest detectors are located 0.75 and 5.78 meters from the source, respectively. Multiple runs were executed with and without the source; each run with the source lasted 120 seconds and each background run lasted 60 seconds.

Bin #	Lower (keV)	Upper (keV)	ISOTOPE(s)
04	123	160	Tc-99m (Technetium 99m) U-235 (Uranium 235)
05	166	203	U-235 (Uranium 235)

TABLE I
Spectral bins for estimating counts for source isotopes.

The dashboard of measurements and counts for an experimental run using a source are shown in Fig. 2. Using the spectra from each detector, the activity levels in two spectral regions associated with potential ^{235}U signatures are estimated as counts at 1 second intervals; the corresponding bin numbers and energy bounds are shown in Table I. They are used as features to train the classifiers for detecting the presence of source. Classifiers are trained and tested using background and source measurements collected over multiple experimental runs under two scenarios:

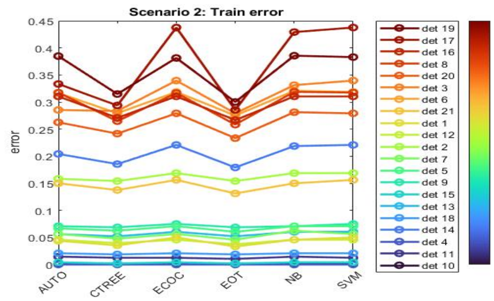
- **S1:** training and testing with 240 counts from 2 background runs and 1 source run; and
- **S2:** training with 480 counts from 4 background and 2 source runs, and testing with 6 background and 4 source runs.

III. Classifiers and Fusers

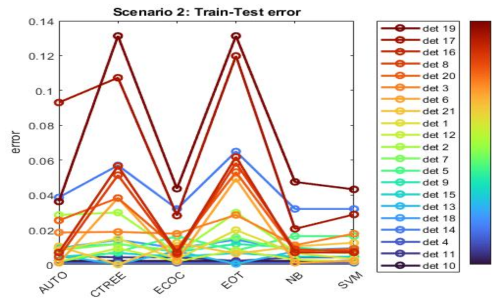
Eight classifiers that represent diverse designs, and six fusers that combine their outputs are considered. The eight classifiers provided by the Matlab ML toolkit are: Auto Tuning and Selection method (AUTO), Classification Trees (CTREE), Error Correcting Output Codes (ECOC), Ensemble of Trees (EOT), k Nearest Neighbors (KNN), Naive Bayes (NB), Neural Network (NN), and Support Vector Machine (SVM). These classifiers are described in [11], [12], and AUTO uses the hyper-parameter searching of individual methods, CTREE, EOT, KNN, NB, and SVM, and chooses one among them based on training data set. These eight different classifiers are chosen to represent their diversity of design, namely smooth and non-smooth, statistical, structural, and hyper parameter tuning and classifier selection methods [1], [4], [6]. The sheer complexity of these classifiers and fusers makes it hard to predict their performance for the current task, and our data sets enable their systematic experimental study.

A. Detector Profiles of Classifiers

The detector profiles of six of eight classifiers, called the *select-six*, are chosen, by excluding the neural network and nearest neighbor methods whose profiles deviated significantly, as described in [12]. The detector profiles of the select-six classifiers in turn reflect the quality of the detector spectrum as well as the detector distance to the source, as shown in Fig. 3.

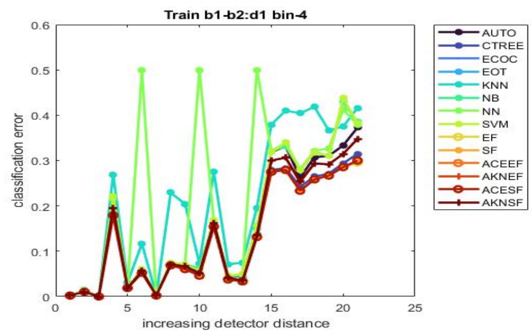


(a) S2: Training

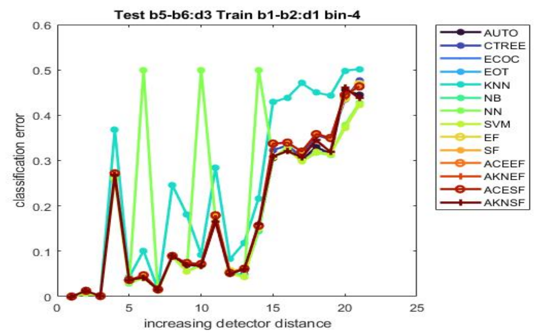


(b) S2: Testing

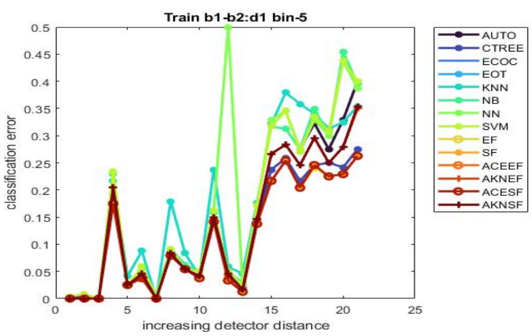
Fig. 4. Classifier profiles of classifiers at different detector distances.



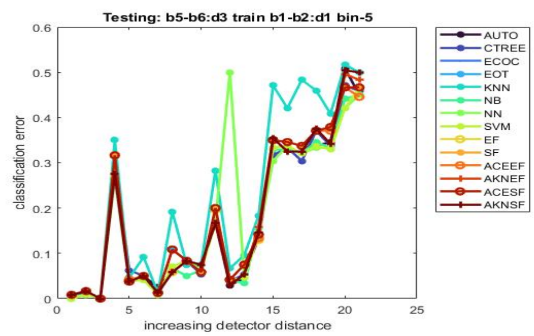
(a) *S1: train - bin 4*



(b) *S1: test - bin 4*



(c) *S1: train - bin 5*



(d) *S1: test - bin 5*

Fig. 5. Distance profiles classifiers and fusers with bins 4 and 5 - they are similar for two bins but different between training and testing for the outer group.

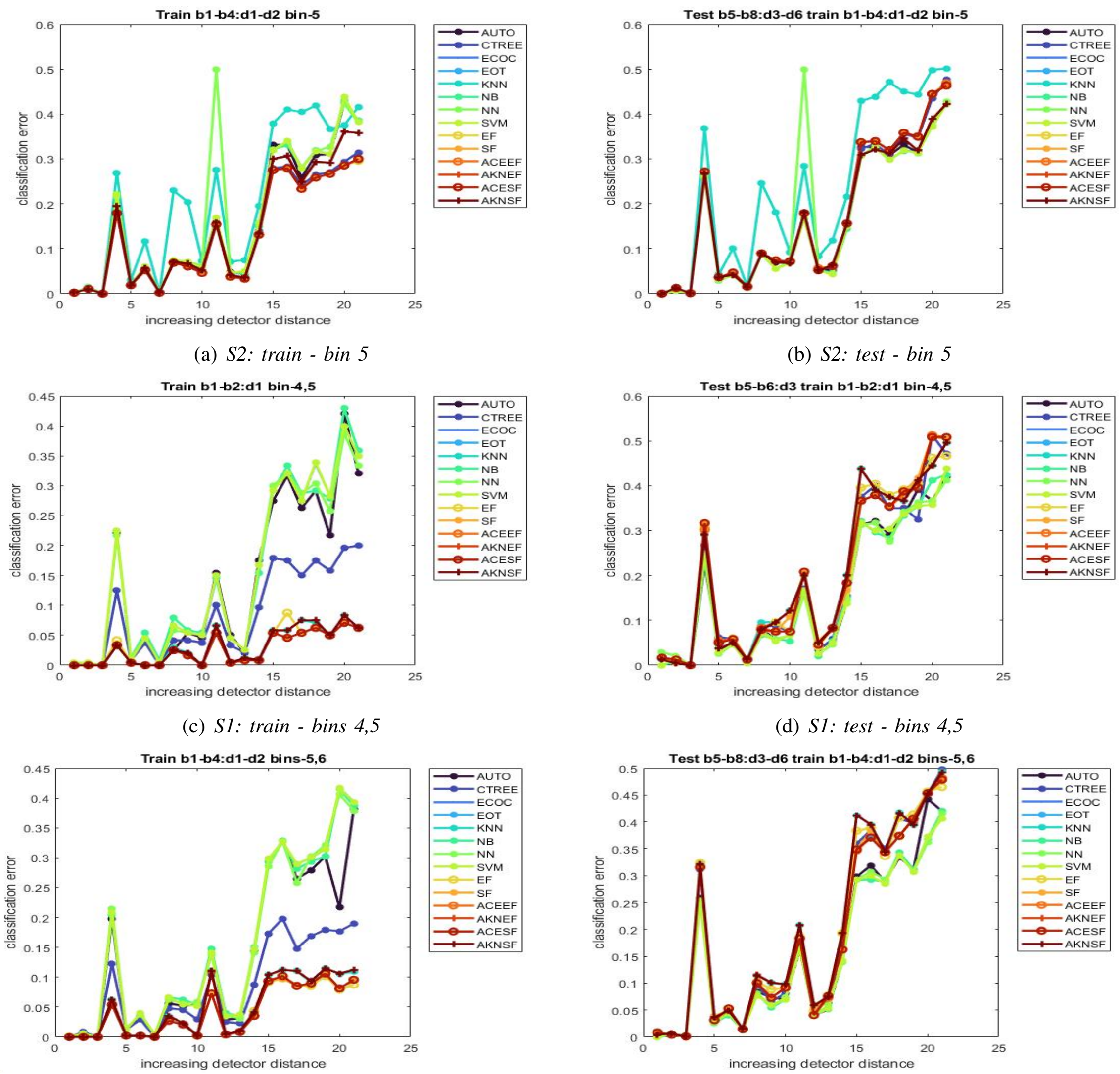
The classification error improves overall as measurements from detectors closer to source are used, as indicated by classifier profiles in Fig. 4. The detectors form three groups of seven each, the inner (75-189 cm), middle (190-317 cm) and outer (359-578 cm), as shown in Fig. 3. They represent an increasing complexity of classification reflected in their range of classification error. The training profiles are well separated, and they become lower as detectors closer to the source are considered, except for outlier detector 14, as shown in Fig. 4(a). This separation does not entirely translate to test profiles, particularly, in the middle and near groups, as shown in Fig. 4(b).

B. Detector Profiles of Classifier-Fusers

The classifiers are fused using two different methods, namely, non-smooth EOT and smooth SVM. Three different sets of classifiers are fused: (a) all eight classifiers are fused using the non-smooth EOT and smooth SVM fusers; these fusers are denoted by EF and SF, respectively; (b) the hyper-parameter and selection classifier AUTO and two non-smooth classifiers CT and EOT are fused using EOT and SVM methods, denoted by ACEEF and ACESF, respectively; and (c) the classifiers AUTO, KNN, NN are fused using EOT and SVM methods, denoted by AKNEF and AKNSF, respectively.

The detector profiles of classifiers and their fusers using two features under scenarios S1 and S2 are summarized in Figs. 5 and 6. The detector profiles using single features, based on bins 4 and 5, indicate similar profiles, as shown in Fig. 5. The fusers achieve lower training error across all three detector groups, but it translates to lower testing error only for the near and middle groups. For the outer group, the classifier-fusers achieve lowest training errors but their test error is higher than those of AUT0 and SVM, which is an indication of over-fitting.

Increasing the data sizes reduces the gap between the training profiles of classifiers and their fusers (Fig. 6(a)), and lowers the gap between training and testing profiles (Fig. 6(b)). Inclusion of an additional feature widens the training gap of



(a) *S2: train - bin 5*

(b) *S2: test - bin 5*

(c) *S1: train - bins 4,5*

(d) *S1: test - bins 4,5*

(e) *S2: train - bins 4,5*

(f) *S2: test - bins 4,5*

Fig. 6. Training and testing detector profiles of classifiers and fusers with larger data sets, two spectral bins and both.

the fusers, and hence the widening of train-test gap, as shown in Figs. 6(c) and (d), respectively. These effects are somewhat lessened when larger data sets are used, as shown in Figs. 6(e) and (f). In both cases of increased features and data sizes, the fusers are subject to more over-fitting (Figs. 6(c) and (e)).

IV. Generalization Equations

To complement the empirical results presented in previous sections, we derive the generalization equations of the classifier-fuser method. A classifier learns a *classification function* $f \in \mathcal{F}$ between an input feature vector $X \in \Re^d$ and an output label $Y \in \{0,1\}$ such that $f(X)$ is an estimate of Y. The training process utilize l-sample $\{(X_1,Y_1),(X_2,Y_2),\ldots,(X_l,Y_l)\}$ to estimate a "suitable" classifier function $\hat{f} \in \mathcal{F}$ by (approximately) minimizing the *empirical misclassification error* given by

$$\hat{I}(f) = \frac{1}{l}\sum_{m=1}^{l}(f(X_m) \oplus Y_m), \tag{1}$$

where $\oplus$ is the Exclusive-OR operation: 0 if $f(X_m) = Y_m$ and 1 otherwise. For individual classifiers, X corresponds to features ($d = 1,2$) and Y is a binary variable, and for classifier-fusers, X belongs to $\{0,1\}^n, n = 3,8$, and Y is binary variable.

In a generic classification problem the feature vector X and output vector $Y \in \{0,1\}$ are distributed jointly accordingly to an unknown distribution $\mathbb{P}_{X,Y}$. The *expected classification error* of a classification function f is

$$I(f) = \int (f(X) \oplus Y)\, d\mathbb{P}_{X,Y}.$$

The *expected best* classifier f^* minimizes $I(.)$ over $\mathcal{F}$, i.e., $I(f^*) = \min_{f\in\mathcal{F}} I(f)$.

The joint distribution $\mathbb{P}_{X,Y}$ of data is complex and is only partially known, since it depends on gamma spectral measurements, sensor errors, and background variations. Consequently, an optimal f^* cannot be computed precisely with probability of one even in principle. Under certain conditions, Vapnik's generalization theory [17] establishes that a "suitable" estimator $\hat{f}$ obtained from independently and identically distributed (iid) training data can ensure

$$\mathbb{P}^{l}_{X,Y}\left[I(\hat{f}) - I(f^*) > \epsilon\right] < \delta(\epsilon, \hat{\epsilon}, l) \quad (2)$$

where $\epsilon, \hat{\epsilon} > 0$, $0 < \delta < 1$, and $\hat{I}(\hat{f}) = \min_{f \in \mathcal{F}} \hat{I}(f) + \hat{\epsilon}$. This condition ensures that "error" of $\hat{f}$ is within ϵ of optimal error (of f^*) with probability $1 - \delta$, *irrespective* of the underlying distribution $\mathbb{P}^{l}_{X,Y}$. Furthermore, the confidence parameter $\delta(\epsilon, \hat{\epsilon}, l)$ approaches 1 as the sample size l approaches infinity. The generalization bound $\delta(\epsilon, \hat{\epsilon}, l)$ applicable to individual classifiers can be derived using various properties of classifier function classes [10].

A. Estimation of Improvement

Consider the fuser class $\mathcal{F}_F$ used in fusing the classifiers $f_a \in \mathcal{F}_a, a \in \mathcal{A}$={AUTO, CTREE, ECOC, EOT, KNN, NB, NN, SVM}. Let f_F denote the classifier function obtained by composing f_a's with the fuser function from $\mathcal{F}_F$. The *improvement* Δ_F of the fused estimate over the best individual classifier is defined as

$$\Delta_F = \min_{a \in A} I(f_a) - I(f_F).$$

Then, if $\mathcal{F}_F$ has the isolation property [9], then $\Delta_F \geq 0$. The best accuracy improvement is given by $\Delta^*_F = \min_{a \in \mathcal{A}} I(f^*_a) - I(f^*_F)$. and its estimate based on samples is given by $\tilde{\Delta}_F = \min_{a \in \mathcal{A}} \hat{I}(f_a) - \hat{I}(f_F)$.

B. Estimation of Confidence

We now show that the estimate $\tilde{\Delta}_F$ is within ϵ of the optimal Δ^*_F with a probability that improves with the training data size l independent of the underlying distribution $\mathbb{P}_{Y,X}$.

Theorem 4.1: Consider that there exists $\delta_b(\epsilon, \hat{\epsilon}_b, l)$ such that based on i.i.d. l-sample, we have

$$\mathbb{P}^{l}_{X,Y}\left[I(\hat{f}_b) - I(f^*_b) > \epsilon\right] < \delta_b(\epsilon, \hat{\epsilon}_b, l). \quad (3)$$

for all classifiers $b = a \in \mathcal{A}$ and classifier-fusers $b = d \in \mathcal{D} = \{EF, SF, ACEEF, ACESF, AKNEF, AKNSF\}$ such that $\delta_b(\epsilon, \hat{\epsilon}_b, l) \to 0$ as $l \to \infty$. Then, the probability that $\tilde{\Delta}_F$ is within ϵ of Δ^*_F is bounded as

$$\mathbb{P}^{l}_{X,Y}\left[|\tilde{\Delta}_F - \Delta^*_F| < \epsilon\right] > 1 - \delta_d(\epsilon/2, \hat{\epsilon}_d, l) - \sum_{a \in \mathcal{A}} \delta_a(\epsilon/(2N_{\mathcal{A}}), \hat{\epsilon}_a, l),$$

for any classifier-fuser $d \in \mathcal{D}$.

Proof: We first note that for $d \in \mathcal{D}$

$$|\tilde{\Delta}_F - \Delta^*_F| \leq |\hat{I}(\hat{f}_d) - I(f^*_d)| + \left|\min_{a \in \mathcal{A}} \hat{I}(\hat{f}_a) - \min_{a \in \mathcal{A}} I(f^*_a)\right|,$$

which establishes that the condition $|\tilde{\Delta}_F - \Delta^*_F| > \epsilon$ implies that at least one term on the right hand side is greater than $\epsilon/2$. We now have

$$|\hat{I}(\hat{f}_d) - I(f^*_d)| \leq |\hat{I}(\hat{f}_d) - I(\hat{f}_d)| + |I(\hat{f}_d) - I(f^*_d)|,$$

which in turn establishes that the condition $|\hat{I}(\hat{f}_d) - I(f^*_d)| > \epsilon/2$ implies that at least one term on the right hand side is greater than $\epsilon/4$. Then, by hypothesis in Eq (3), both conditions are simultaneously satisfied with probability at most $\delta_d(\epsilon/4, \hat{\epsilon}_d, l)$. Similarly, we have

$$\left|\min_{a \in \mathcal{A}} \hat{I}(\hat{f}_a) - \min_{a \in \mathcal{A}} I(f^*_a)\right| \leq \left|\min_{a \in \mathcal{A}} \hat{I}(\hat{f}_a) - \min_{a \in \mathcal{A}} I(\hat{f}_a)\right| + \left|\min_{a \in \mathcal{A}} I(\hat{f}_a) - \min_{a \in \mathcal{A}} I(f^*_a)\right|,$$

which in turn establishes that the condition $\left|\min_{a \in \mathcal{A}} \hat{I}(\hat{f}_a) - \min_{a \in \mathcal{A}} I(f^*_a)\right| > \epsilon/2$ implies that at least one term on the right hand side is greater than $\epsilon/4$. Then, we consider the two upper bounds

$$\left|\min_{a \in \mathcal{A}} \hat{I}(\hat{f}_a) - \min_{a \in \mathcal{A}} I(\hat{f}_a)\right| \leq \sum_{a \in \mathcal{A}} \left|\hat{I}(\hat{f}_a) - I(\hat{f}_a)\right|$$

$$\left|\min_{a \in \mathcal{A}} I(\hat{f}_a) - \min_{a \in \mathcal{A}} I(f^*_a)\right| \leq \sum_{a \in \mathcal{A}} \left|I(\hat{f}_a) - I(f^*_a)\right|.$$

In each case, the condition that left hand side is larger than $\epsilon/2$ implies at least one of the terms under the summation is greater $\epsilon/(2N_{\mathcal{A}})$. Under the hypothesis of this theorem in Eq (3), both conditions are satisfied with probability at most

$$\sum_{a \in \mathcal{A}} \delta_a(\epsilon/(2N_{\mathcal{A}}), \hat{\epsilon}_a, l).$$

By combining the above terms together, we have

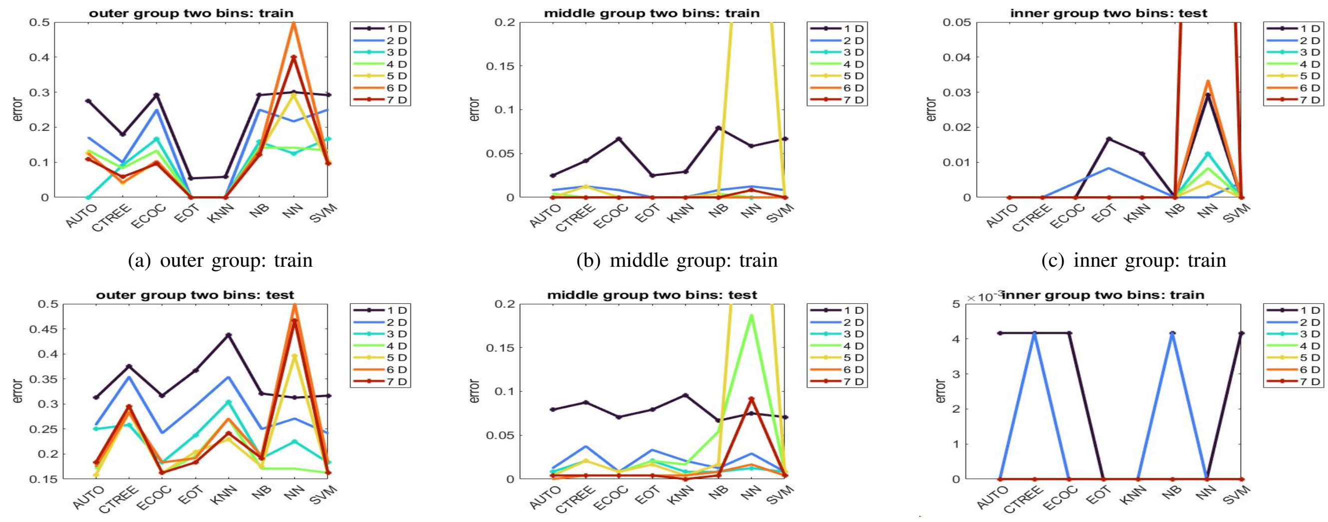
$$\mathbb{P}^{l}_{X,Y}\left[|\tilde{\Delta}_F - \Delta^*_F| > \epsilon\right] < \delta_d(\epsilon/2, \hat{\epsilon}_d, l) + \sum_{a \in \mathcal{A}} \delta_a(\epsilon/(2N_{\mathcal{A}}), \hat{\epsilon}_a, l),$$

which proves the theorem. □

The confidence bound is expressed in terms of the *precision* parameter ϵ and the *confidence* parameter $[1 - \delta_d(\epsilon/2, \hat{\epsilon}_d, l) - \sum_{a \in \mathcal{A}} \delta_a(\epsilon/(2N_{\mathcal{A}}), \hat{\epsilon}_a, l)]$, which approaches 1 with increasing number of measurements l, thereby indicating the effectiveness of the classifier fuser. However, this result is asymptotic and is not fine enough to provide inferences for finite (small) samples. For a fixed l, the confidence on the precision of $\tilde{\Delta}$ is lower in proportion to the number of fused classifiers and also their parameters; it is reflected in the higher generalization error of classifier-fusers despite their lower training in the previous section.

V. Detector Fusers

The features from multiple detectors are fused within each group, starting with the one nearest to the source and successively adding detectors in the increasing order of their distance. There is an overall lowering of the classifier profiles as more detectors are fused in the outer group, both for training and testing as shown in Fig. 3 (a) and (d), respectively. The error reduction is least for the inner group, since the single detector classifiers already achieve low errors, as shown in Fig. 3 (c) and (f), respectively. The lowering of profiles is very instructive in the middle group, where the addition of the second detector resulted in visible lowering of the profile, but further addition of detectors has less effect. Overall, the addition of detectors effectively increases the capture area, albeit distributed, of NaI material, which is exposed to



(a) outer group: train (b) middle group: train (c) inner group: train

(d) outer group: test (e) middle group: test (f) inner group: test

Fig. 7. Classifier error profiles of fusers with 1, 2, ..., 7 detectors in outer, middle and inner groups.

more gamma radiation. But, the performance saturates once it reaches around that of the nearest detector adjusted for distance, wherein further addition of detectors does not lead to much improved performance.

VI. Conclusions

By using controlled measurements, three types of information fusion strategies are studied for detecting low-level ^{235}U radiation sources, using features obtained from gamma spectra of NaI detectors. This study provides two main qualitative insights into this classification task. First, the fusion of detectors leads to an overall improved classification performance, explained by increased effective NaI capture area. Second, the over-fitting by classifiers and classifier-fusers calls for methods supported by the underlying physics, analytical and statistical models, rather than simply relying on complex machine learning methods that perform well on training data.

Future directions include developing physics, data and statistical basis for the generalization performance of various classifiers, and quantitative measures to assess and compare the profiles to complement qualitative approach of this paper. It would be interesting to see if over-fitting of the classifiers is specific to the codes used, by testing other codes such as scikitpy and R. It would be of future interest to study similar aspects for other low level radiation sources.

References

[1] E. Alpaydin. *Introduction to Machine Learning*. MIT Press, 2020. fourth edition.

[2] S. M. Brennan, A. M. Mielke, and D. C. Torney. Radiation detection with distributed sensor networks. *IEEE Computer*, pages 57–59, August 2004.

[3] C. Drummond. An incremental machine learning algorithm for nuclear forensics. In E. Bagheri and J. C. Cheung, editors, *Advances in Artificial Intelligence*, pages 194–207, Cham, 2018. Springer International Publishing.

[4] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley and Sons, Inc., 2001. Second Edition.

[5] M. Girard, A. Hagen, I. Schwerdt, M. Gaumer, L. McDonald, N. Hodas, and E. Jurrus. Uranium oxide synthetic pathway discernment through unsupervised morphological analysis. *Journal of Nuclear Materials*, 552, 2021.

[6] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, 2001.

[7] M. J. Kristo and S. J. Tumey. The state of nuclear forensics. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 294:656–661, 2013. Proceedings of the Twelfth International Conference on Accelerator Mass Spectrometry, Wellington, New Zealand, 20-25 March 2011.

[8] J. R. Lewis, A. Zhang, and C. M. Anderson-Cook. Comparing multiple statistical methods for inverse prediction in nuclear forensics applications. *Chemometrics and Intelligent Laboratory Systems*, 175:116–129, 2018.

[9] N. S. V. Rao. Measurement-based statistical fusion methods for distributed sensor networks. In S. S. Iyengar and R. R. Brooks, editors, *Distributed Sensor Networks*. Chapman and Hall/CRC Publishers, 2011. 2nd Edition.

[10] N. S. V. Rao. Generalization equations for machine learners based on physical and abstract laws. In *IEEE International Conference on Multi-Sensor Fusion and Integration*, 2021.

[11] N. S. V. Rao, C. Greulich, M. P. Dion, J. Hite, K. J. Dayman, A. D. Nicholson, D. E. Archer, M. J. Willis, J. M. Ghawaly, I. G. R. D. Hunley, and J. Johnson. Isotope ratio features for classification of dissolution events using effluents measurements. In *Institute of Nuclear Materials Management Annual Meeting*, 2021.

[12] N. S. V. Rao, D. Hooper, and J. Ladd-Lively. Study of classifiers for u-235 source signatures using gamma spectral measurements. In *Institute of Nuclear Materials Management Annual Meeting*, 2022.

[13] N. S. V. Rao, S. Sen, M. L. Berry, C. Q. Wu, K. M. Grieme, R. R. Brooks, and G. Cordone. Datasets for radiation network algorithm development and testing. In *2016 IEEE Nuclear Science Symposium*, 2016.

[14] N. S. V. Rao, S. Sen, N. J. Prins, D. A. Cooper, R. J. Ledoux, J. B. Costales, K. Kamieniecki, S. E. Korbly, J. K. Thompson, J. Batcheler, R. R. Brooks, and C. Q. Wu. Network algorithms for detection of radiation sources. *Nuclear Instruments and Methods Phy. Res. A*, 784:326–331, June 2015.

[15] B. Ristic, M. Morelande, and A. Gunatilaka. Information driven search for point sources of gamma radiation. *ACM Signal Processing*, 90(4):1225–1239, 2010.

[16] D. L. Stephens and A. J. Peurrung. Detection of moving radioactive sources using sensor networks. *IEEE Transactions on Nuclear Science*, 51(5):2273–2278, 2004.

[17] V. N. Vapnik. *Statistical Learning Theory*. John-Wiley and Sons, New York, 1998.

[18] J. R. Vavrek, D. Hellfeld, M. S. Bandstra, V. Negut, K. Meehan, W. J. Vanderlip, J. W. Cates, R. Pavlovsky, B. J. Quiter, R. J. Cooper, and T. H. Y. Joshi. Reconstructing the position and intensity of multiple gamma-ray point sources with a sparse parametric algorithm. *IEEE Transactions on Nuclear Science*, 67(11):2421–2430, 2020.