Broad Area Search and Detection of Surface-to-Air Missile Sites Using Spatial Fusion of Component Object Detections From Deep Neural Networks

Alan B. Cannaday II, *Student Member, IEEE*, Curt H. Davis, *Fellow, IEEE*, Grant J. Scott, *Senior Member, IEEE*, Blake Ruprecht, *Student Member, IEEE*, and Derek T. Anderson, *Senior Member, IEEE*

***Abstract*—Here, we demonstrate how deep neural network (DNN) detections of multiple constitutive or component objects that are part of a larger, more complex, and encompassing feature can be spatially fused to improve the search, detection, and retrieval (ranking) of the larger complex feature. First, scores computed from a spatial clustering algorithm are normalized to a reference space so that they are independent of image resolution and DNN input chip size. Then, multiscale DNN detections from various component objects are fused to improve the detection and retrieval of DNN detections of a larger complex feature. We demonstrate the utility of this approach for broad area search and detection of surface-to-air missile (SAM) sites that have a very low occurrence rate (only 16 sites) over a ∼90 000 km^2 study area in SE China. The results demonstrate that spatial fusion of multiscale component-object DNN detections can reduce the detection error rate of *SAM Sites* by >85% while still maintaining a 100% recall. The novel spatial fusion approach demonstrated here can be easily extended to a wide variety of other challenging object search and detection problems in large-scale remote sensing image datasets.**

***Index Terms*—Broad area search, data fusion, deep neural networks (DNN), information retrieval, object detection, spatial clustering.**

I. Introduction

WITHIN the last five years, deep neural networks (DNN) have shown through extensive experimental validation to deliver outstanding performance for object detection/recognition in a variety of benchmark high-resolution remote sensing image datasets [1]–[7]. Methods such as you only look once [8], region-based CNN (R-CNN) [9], and derivations thereof [10]–[15] have all shown promising results for a variety of object detection applications in remote sensing imagery.

The demonstrated ability of DNNs to automatically detect a wide variety of man-made objects with very high accuracy has tremendous potential to assist human analysts in labor-intensive visual searches for objects of interest in high-resolution imagery over large areas of the Earth's surface. However, the vast majority of published studies for DNN object detection in remote sensing imagery have focused on development of new deep learning algorithms/methods and/or comparative testing/evaluation of these methods on benchmark datasets (both public and private).

As noted by Xin *et al.* [16], comparatively fewer studies have attempted to apply promising DNN methods to demonstrate efficacy and/or further develop these new methods via applications to large-scale or broad area remote sensing image datasets, e.g., [17]–[19]. Since "large-scale" or "broad area" are subjective descriptors; here, we define these to be applications where the algorithm is applied to validation image datasets, i.e., excluding training data, covering an area greater than 1,000 km^2.

Furthermore, even DNN detectors that demonstrate exceptionally high accuracy (e.g., 99%) on benchmark testing datasets will still generate a tremendous number of errors when applied to large-scale/broad area remote-sensing image datasets. For example, a DNN detector with 99% average accuracy, chip size of 128 × 128 pixels, and a chip scan overlap of 50% will generate 88000 errors when applied to a 0.5 m ground sample distance (GSD) image dataset covering an area of interest (AOI) of 10 000 km^2 (e.g., 1° × 1° cell).

If post DNN detection results are intended to be reviewed by human analysts in machine-assisted analytic workflows, then large numbers of detection errors can quickly lead to "error fatigue" and a corresponding negative end-user perception of machine-assisted workflows. Thus, it is important to develop methods to reduce error rates resulting from application of DNN detectors to large-scale/broad area remote-sensing image datasets to improve machine-assisted analytic workflows.

In this study, we develop a new framework for spatially fusing multiscale detections from a variety of component objects to improve the detection and retrieval of a larger complex feature. A key aspect of this framework is the development of a spatial clustering algorithm that generates normalized per-object cluster scores to facilitate the spatial fusion of the component objects detected at variable image resolutions and spatial extents. We demonstrate the efficacy of this approach in a broad area search and detection application of surface-to-air missile (SAM) sites over a large search area where >85% error reduction is achieved

Manuscript received March 20, 2020; revised June 3, 2020 and July 20, 2020; accepted July 29, 2020. Date of publication August 11, 2020; date of current version August 26, 2020. *(Corresponding author: Alan B. Cannaday II.)*

Alan B. Cannaday II, Curt H. Davis, and Grant J. Scott are with the Center for Geospatial Intelligence, University of Missouri, Columbia, MO 65211 USA (e-mail: abcannaday@mail.missouri.edu; DavisCH@missouri.edu; GrantScott@missouri.edu,).

Blake Ruprecht and Derek T. Anderson are with the Mizzou Information and Data Fusion Laboratory, University of Missouri, Columbia, MO 65211 USA (e-mail: bcrf53@mail.missouri.edu; andersondt@missouri.edu).

Digital Object Identifier 10.1109/JSTARS.2020.3015662

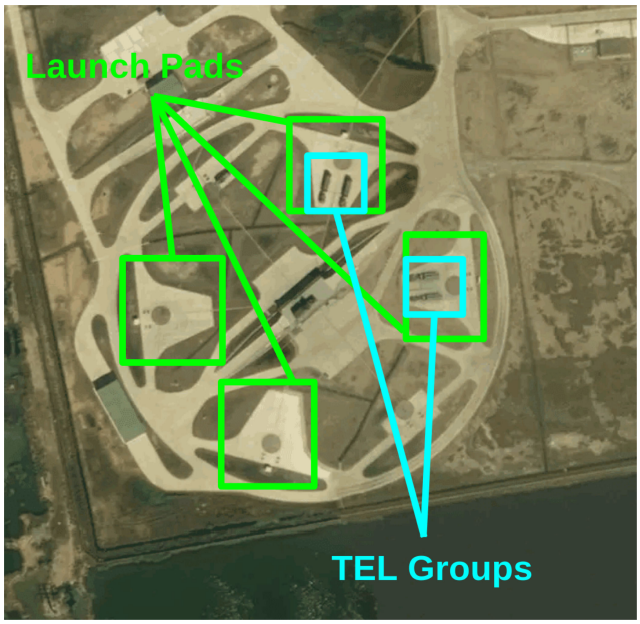

Fig. 1. Example *SAM Site* with smaller-scale *LP* and *TEL Group* component objects.

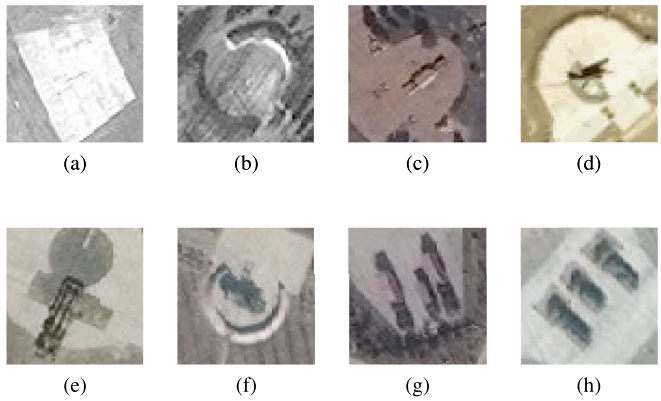

Fig. 2. Samples of *SAM Site* component objects used in this study. (a) Empty LP. (b) Empty LP. (c) Missile. (d) Missile. (e) TEL. (f) TEL. (g) TEL Group. (h) TEL Group.

TABLE I
SUMMARY OF CURATED TRAINING DATA

Object Class	*SAM Sites*	*Launch Pads*	*Missiles*	*TELs*	*TEL Groups*
TP	101	3910[1]	1976	2733	1179
TN	404	3696	2624	2272	1054
Combo TP	n/a	9798[2]	as above	as above	as above
Combo TN	n/a	8512	6530	10,078	5762

1: Empty.
2: Includes those with colocated *missiles*, *TELs*, and *TEL Groups*.

for a 100% recall. This new framework can be easily adapted or extended to a variety of other challenging object search and detection problems in large-scale remote sensing image datasets.

II. Study Area and Source Data

This study builds upon Marcum *et al.* [19] where broad area search and detection of *SAM Sites* (Fig. 1) was demonstrated over a ~90 000 km^2 study AOI along the SE coast of China. Key results from the prior study were as follows.

1) A machine-assisted approach was used to reduce the original AOI search area by 660 X to only ~135 km^2.
2) The average machine-assisted search time for ~2100 candidate *SAM Site* locations was ~42 min which was 81 X faster than a traditional human visual search.

While Marcum *et al.* used a single binary DNN detector to locate candidate *SAM Sites*; here, we explore the benefit of fusing multiscale DNN detectors of smaller component objects to improve the detection of the larger encompassing *SAM Site* features.

First, a binary *SAM Site* DNN detector was trained using a slightly enhanced version of the curated *SAM Sites* training data in China from [19]. To ensure blind scanning, only 101 *SAM Sites* lying outside the SE China AOI were used to train the DNN. While [19] used a 227 × 227 pixel chip size at 1 m GSD to train a ResNet-101 [20] DNN; here, we used a 299 × 299 pixel chip size at 1 m GSD for training a neural architecture search network [21] (NASNet) architecture.

As in [19], negative training chip samples were selected using a 5-km offset in the four cardinal directions (i.e., N/S/E/W) for each *SAM Site*. The SE China AOI has 16 known *SAM Sites,* which includes two newer *SAM Sites* found in the previous study [19].

We next developed binary DNN detectors for four different *SAM Site* component objects: *Launch Pads (LPs)*, *missiles*, *transporter erector launchers* (*TELs*), and *TEL Groups* (two or more ~colocated *TELs*) (Fig. 2). Component binary DNN detectors were trained using curated data at 0.5 m GSD from China *SAM Sites* outside the AOI. We first created negative training samples for each component using nearby image chips (similar land cover context), but outside the known spatial extent of a *SAM Site*. This produced a ~1:1 ratio of negative to positive component training samples (Table I).

In addition, we created a second training dataset using all four components to train a combined *LP* detector (empty and nonempty) knowing that the other components (e.g., *missiles*, *TELs*, etc.) are generally colocated with *LPs*. We then developed a second set of component detectors for the *missile*, *TEL*, and *TEL Group* object classes by combining negative training data from the other components and then randomly paring down the data to produce a ~4:1 ratio of negative to positive samples (Table I). For the *missile* component, samples from empty *LPs*, *TEL*, and *TEL Group* and their negatives were added. However, only samples from empty *LPs* and *missiles* were added to the negatives for *TELs* and *TEL Groups* to reduce confusion between these two components.

Different chip sizes were used for the training samples based on known object sizes. A 128 × 128 pixel chip size was used for detecting both empty and combined *LPs* and *TEL Groups*. While a 64 × 64 pixel chip size was used for *Missiles* and *TELs*. Counts for all training data are provided in Table I and these

TABLE II
SUMMARY OF *SAM SITE* DNN DETECTOR PERFORMANCE FROM FIVE-FOLD CROSS-VALIDATION. METRICS SHOWN ARE *TPR*, *TNR*, *ACC*, AND *AUC*

DNN	*TPR* (%)	*TNR* (%)	*ACC* (%)	*AUC* (%)
ResNet-101	94.1%	98.8%	96.4%	99.4%
Xception	98.0%	98.3%	98.1%	99.9%
NASNet	99.0%	99.8%	99.4%	99.995%
ProxylessNAS	95.0%	100.0%	97.5%	99.2%
EfficientNet-B4[1]	91.1%	99.8%	95.4%	99.8%

1: Only results from the EfficientNet model with the highest *AUC* are shown.

only include component samples outside the SE China AOI to ensure blind scanning.

III. DATA PROCESSING

A. Training Data Augmentation

Augmentation strategies from [19] were used to train the *SAM Site* DNN and all component DNNs to improve detector performance. A $144X$ augmentation was used for all fivefold validation experiments while a $9504X$ augmentation was used for the final *SAM Site* DNN used in the AOI scanning. To save computing time, augmentations were reduced for training the component DNNs due to the much larger sample sizes. These changes included using RGB samples only, reducing the number of rotations, using a single jitter distance, and removing the contrast augmentation. Most of the final component DNNs were trained with $648X$ augmentations, except the combined *LP* DNN used a $216X$ augmentation.

B. Scanning and Spatial Clustering

We completed fivefold cross-validation experiments for several modern DNN architectures to evaluate their performance for *SAM Site* detection. The modern DNN architectures we evaluated were NASNet [21], Xception [22], ProxylessNAS [23], and all seven EfficientNet [24] models. *SAM Site* detection results from these modern DNNs are compared to the ResNet-101 DNN results published in the Marcum *et al.* [19] study (Table II).

The results in Table II show that the NASNet DNN outperformed all the other DNNs for *SAM Site* detection as measured by *true negative rate (TPR)*, *average accuracy (ACC)*, and *area under the ROC curve (AUC)*. In addition to training the *SAM Site* DNN detector, the NASNet DNN was used for training all component DNN detectors used throughout the rest of this study. Training for all the NASNet DNNs utilized transfer learning from ImageNet [25], Adam [26] for optimization, and cross entropy for the objective function.

As in [19], images used for broad area search for *SAM Sites* in the SE China AOI were comprised of ∼66 K 1280 × 1280 pixel tiles at 1 m GSD with 10% overlap between tiles. Individual tiles were scanned by generating ∼19.7 M image chips with 75% overlap (25% stride) that were then input to the NASNet models. This produced a raw detection field, F, of softmax outputs from the DNN. After thresholding F at $\alpha = 0.9$, a greatly reduced detection field, F^{α}, is then used to produce an amplified spatial detection field, δ. The δ is used to weight a spatial clustering of F^{α} to produce mode clusters, F', within a 300-m aperture radius, R (see [19]). Cluster locations were then rank-ordered by summing the scores of all detections within a mode cluster to generate an initial set of "candidate" *SAM Sites*.

A new 1280 × 1280 pixel tile at 0.5 m GSD centered on each candidate *SAM Site*'s cluster location was then used for all component DNN scans. Component scanning outputs were also spatially clustered to generate locations and cluster scores for each component object. An aperture radius of $R = 32$ m was used since this is approximately half the typical distance between *SAM Site* LPs in China. An alpha cut of $\alpha = 0.99$ was used to generate distinct cluster locations for a given component relative to neighboring components that were present at each candidate *SAM Site*. In this study, we simply used *a priori* knowledge for our selection of *R*. However, we recognize that for other objects and/or applications the appropriate selection of *R* may need to be incorporated in the technical approach, as it can be sensitive to scanning stride and target object colocation separation (e.g., vehicles parked next to each other).

Likewise, in order to determine thresholds used for the decision-theoretic approach (DTA) described in Section III-E, a training set of 1280 × 1280 pixel pseudocandidate tiles were generated that were centered about the known *SAM Sites* outside the SE China AOI along with corresponding offset tile negatives. The same scans and processes performed for candidate tiles within the SE China AOI (described abovementioned), were used for the pseudocandidate training dataset.

C. Cluster Score Normalization and Truncation

Cluster scores from one object class to another are not necessarily comparable since they can result from objects with different physical sizes, corresponding *R* values, and scanning strides. In addition, results generated from image tile scans with different DNN input chip size and/or GSD will have a variable spatial density. Since we wish to spatially fuse, and potentially weight, the output from various component DNN detectors, the cluster scores must be normalized to bring both the *SAM Site* and component detection clusters into a common reference space. Here, we use raw detection field density, i.e., the number of raw detections per unit area, as the means to achieve a common reference space prior to spatially fusing the cluster scores from the candidate *SAM Sites* and their associated component detection clusters.

1) Normalization for a Single Detection Location: The amplified spatial detection field, δ, contains an intersected volume, δ_n, for each raw detection, n, in F^{α}. δ_n is calculated as the weighted sum of scores of each n with its neighboring raw detections, p. The weight is determined by the distance-decay function $s(p) = exp^{(-d/R)}$, where $d = \text{haversine}(p, n) < R$ and is 0 otherwise. An approximate maximum intersection volume for a single raw detection can be calculated by integrating the truncated distance-decay function around a raw detection location. As mentioned above, R and d are normalized using raw detection field density. Let $R' = R/\text{stride}$ and $d' = d/\text{stride}$, where *stride* is the image chip's scanning stride distance in meters, so that $s'(p) = \exp^{(-d'/R')}$. Let s represent the height or the dimension of magnitude for F, then the approximate

Procedure 1: Object Detection Cluster Ranking with Normalized Scores and Optional Penalty.

Input: Alpha-cut F^{α}, Mode-Cluster F'
Output: Ranked Clusters C_i, where $C_i < C_{i+1}$
begin
  $i := 0$
  while $F' = \emptyset$ **do**
    $p := pop(F')$
    $C_i := p$ // Init. C_i with chip, p
    $N^{\alpha} := NN(p, F^{\alpha}, R)$
    $N := NN(p, F,' R)$
    for all $n \in N(p)$ **do**
      $C_i := \{C_i, n\}$ $F'.remove(n)$
      if $n \subset N^{\alpha}(p)$ **then**
        $n_{weight} = 1$
      else
        $n_{weight} =$ penalty // 0 if no penalty
      end if
    end for
    $C_i.score = (C_{norm})^{-1} \cdot \sum_{n \in C_i} n_{weight} \cdot \delta_n$
    i++
  end while
  $\{C_i\} :=$ **sort** $C_{\forall i}$ by score, descending
end

max intersection volume for a single raw detection can then be calculated as follows:

$$n_{\text{volume}} = \pi \cdot (R')^2 \left(\left(\int_{1/e}^{1} log^2(1/s) ds \right) + (1/e) \right)$$
$$= \pi \cdot (R')^2 ((2 - 5/e) + (1/e))$$
$$= \pi \cdot (R')^2 (2 - 4/e) \quad (1)$$

2) Cluster Score Weighting/Truncation: The cluster score, C_{score}, should also be limited for normalization. In the previous algorithm from [19], the number of raw detections in a cluster, C, was virtually unbounded. As a result, raw detections that were a large distance away from the cluster location can potentially contribute to C_{score}. We have often observed that these far away detections are commonly false positives. In order to minimize their impact, it can be beneficial to weight each raw detection within a cluster based on the raw detection distance relative to the cluster location. Alternately, truncation can be implemented by assigning each raw detection within a haversine distance *R* a weight of 1 and outside of *R* a weight of 0. The C_{score} is then a weighted sum of the raw detection inference scores with their respective weights (Procedure 1). In Section III-D, we discuss the possibility of applying a negative penalty weight to all raw detections with a haversine distance greater than R.

3) Approximate Max Cluster Score: The number of raw detections within R can be seen as a Gauss circle problem. Thus, the max number of raw detections within the aperture area surrounding a cluster location can be approximated in terms of raw detection field density as follows:

$$n_{\text{max\_p}} = \pi \cdot (R')^2 \quad (2)$$

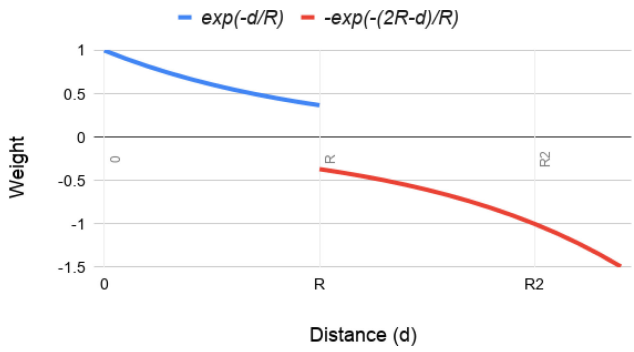

Fig. 3. Distance-decay functions used for calculating local clustering scores. The function $\exp(-d/R)$ (blue) is used as a weight when summing raw detections within distance R. The function $-\exp(-(2R-d)/R)$ (red) is used to calculate an exponential penalty weight for raw detections outside R.

TABLE III
SAMPLE THRESHOLDS CALCULATED BY DTA

Feature Type	*Empty LPs*	*Combo LPs*	*Missiles*	*TELs*	*TEL Groups*
Cluster Count	2	1	1	3	1
Raw Count	5	4	4	15	1
Raw Max	1.00000	0.99954	1.00000	1.00000	0.58236
Including Component Negatives					
Cluster Count	n/a	n/a	1	1	1
Raw Count	n/a	n/a	2	2	1
Raw Max	n/a	n/a	0.99989	0.98450	0.60315

Using (1) and (2), a normalizing cluster factor can be calculated as follows:

$$C_{\text{norm}} = n_{\text{volume}} \cdot n_{\text{max\_p}}$$
$$= \pi^2 \cdot (R')^4 (2 - 4/e) \quad (3)$$

D. Over-Detection Penalty

In previous work, we have observed FP hotspots, i.e., large numbers of spatially cooccurring false positive detections. In order to mitigate this potential problem, a penalty can be applied when computing C_{score}. As mentioned in Section III-C2, instead of using a weight of 0 when $d > R$, a negative weight can be applied. We explored two types of penalty assignments. The first used a flat weight of –1. The second is similar to the distance-decay function; however, the sign was changed to negative and increases in value exponentially as d increases (Fig. 3). The penalty is calculated using the following formula: $s(p) = -\exp(-(2R - d)/R)$.

E. DTA for Optimization

In order to make discrete decisions, we used the DTA [28] advocated by Lewis [29] that computes thresholds based on the optimal prediction of a model to obtain the highest expected F-measure. In this study, decision thresholds were selected based on the optimization of the *F1* score from features extracted from the pseudocandidate training dataset. (Fig. 4). Optimal *F1* score thresholds were determined through empirical analysis and selected examples are provided (Table III).

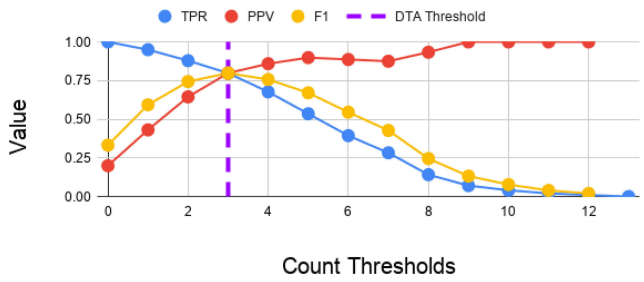

Fig. 4. TPR or recall positive predictive value (PPV or precision), and F1 score versus the threshold for the cluster count of TEL cluster centers within 150 m of a candidate SAM Site location. In this example, the value 3 was used for the final threshold as shown in Table III.

TABLE IV
NASNET FIVEFOLD CROSS-VALIDATION RESULTS FOR DNN MODELS OF *SAM SITES* AND EACH COMPONENT, INCLUDING COMPONENT MODELS WITH NEGATIVE COMPONENT DATA. METRICS SHOWN ARE *TPR*, *TNR*, *F1* SCORE, AND STANDARD DEVIATION (*SD*)

Object Class	*TPR* (%)	*TNR* (%)	*F1* (%)	*SD*
SAM Site	99.00	99.75	99.39	1.06
Empty LPs	99.80	99.70	99.65	0.2
Combo LPs	99.74	99.74	99.74	0.15
Missiles	99.8	99.46	99.63	0.24
TELs	99.72	99.38	99.55	0.32
TEL Groups	99.41	99.43	99.42	0.21
Including Component Negatives				
Missiles	97.42	99.66	98.52	0.72
TELs	97.51	99.37	98.42	0.5
TEL Groups	96.78	99.6	98.15	1.1

TABLE V
SPATIAL CLUSTERING RESULTS FROM DNN SCANNING OF THE SE CHINA AOI FOR CANDIDATE *SAM SITES*. GIVEN VALUES ARE PRECLUSTER COUNTS OVER α-CUT THRESHOLD (F^{α}), POSTCLUSTER COUNTS, AND AVERAGE TRUE POSITIVE (*TP*) CLUSTER RANK

DNN Architecture & Post-Procsessing	F^{α} Count	C Count	AVG *TP* Cluster Rank
ResNet-101 [19]	93,000	2100	181.9
NASNet	2079	354	62.8
NASNet w/ norm	2079	354	62.8
NASNet w/ norm and penalty	2079	354	62.8

IV. FIVEFOLD EXPERIMENT RESULTS

Fivefold cross-validation experiments were performed for the training datasets in Table I. The results provided in Table IV show an average *F1* score $> 99\%$ for the baseline dataset and $> 98\%$ for the dataset with component negatives. The decrease in *F1* score for the DNNs with component negatives was anticipated given the inclusion of objects in the negative training data that were visually similar to the component object that a given DNN was trained to detect.

NASNet significantly outperformed ResNet-101 for scanning the SE China AOI for *SAM Sites* (Table V). This is consistent with the cross-validation results given in Table IV. NASNet had $\sim 44X$ fewer *SAM Site* candidate locations after the 0.9 alpha-cut (Section III-B). Furthermore, while both DNNs correctly located all 16 known *SAM Sites* [e.g., true positive (TPs)] in the SE China AOI, NASNet had $6X$ fewer candidates compared to ResNet-101 while the average TP cluster rank (Table V) was also $\sim 3X$ lower.

V. DECISION-LEVEL COMPONENT METRIC FUSION

This section describes the feature selection and fusion techniques used to reduce the number of candidate *SAM Sites* that could then be presented for human review in machine-assisted analytic workflows. An overview of the processing flow is provided in Fig. 5.

A. Component Features

Five different feature types were used in [27] for decision-level fusion of component objects for improving the final detection of construction sites. Here, we tested feature types that used the *F1* score optimization from [27] and represent the first three feature types listed below. We used the normalized cluster scores from the spatial clustering as an additional feature type. To maintain consistency between techniques employed in this study, only inference responses within a 150 m radius of the candidate *SAM Site* location were used. The feature types that were evaluated were as follows.

1) Maximum raw inference detection response (confidence value) for each component.
2) Count of raw inference detections for each component retained within the reduced field (F^{α}).
3) Count of clusters produced for each component.
4) Sum of normalized cluster scores for each component.

B. Decision-Level Fusion Techniques

Baseline results for the candidate *SAM Site* locations are first computed using only the spatial cluster outputs of the NASNet *SAM Site* detector. We then tested how each individual component would perform using the various feature types. *SAM Site* cluster scores were excluded because the pseudocandidate training dataset was not generated through scanning and clustering. Consequently, some of the pseudocandidates would have no cluster within a sufficient radius of the *SAM Site* center location.

Three data fusion techniques were tested.

1) *Decision Tree*: A simple decision tree (see [27]) was used to combine the decisions generated for each component using DTA. However, unlike [27], this study does not use an alpha-cut threshold since this was part of the spatial clustering algorithm. Therefore, the decision tree is simplified to a digital logic OR gate with the DTA decisions as binary inputs.
2) *Multilayer Perceptron* (MLP): A feature vector was created for each candidate *SAM Site* location and used as input for training and validation. The MLP architecture consisted of two fully connected hidden layers of 100 nodes. We also tested normalization and feature bounds before being used as input based on the thresholds from DTA optimization (Section III-E).

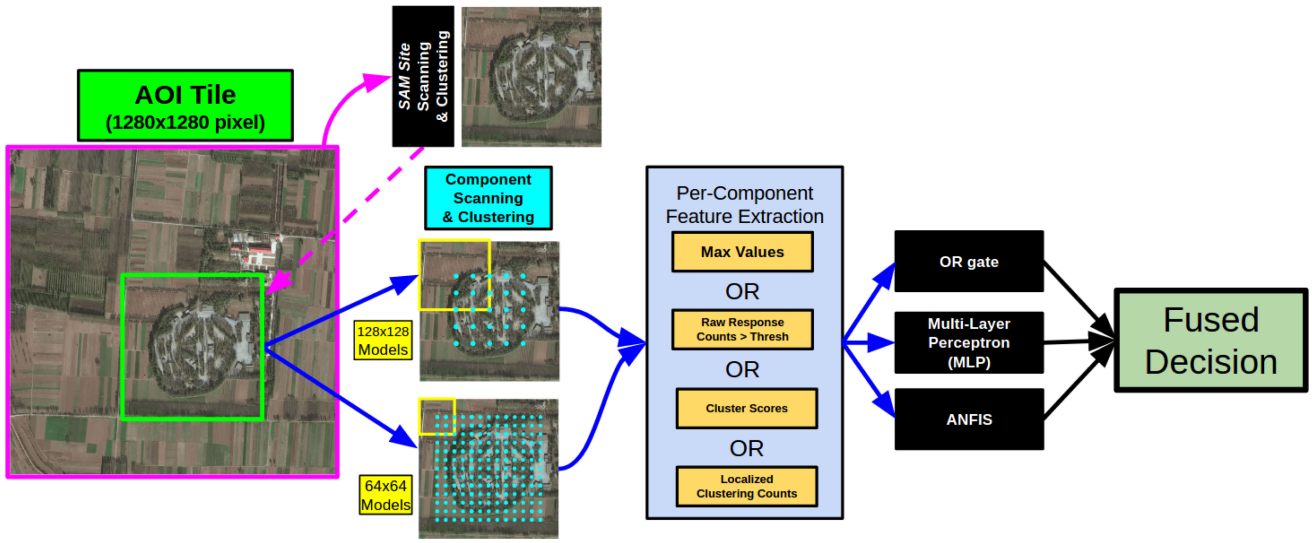

Fig. 5. Processing flowchart for decision-level fusion of multiple component object detections.

3) *ANFIS*: A first-order Takagi–Sugeno–Kang adaptive neuro-fuzzy inference system (ANFIS) [30]–[32] was utilized. The goal is to explore a neural encoding and subsequent optimization of expert knowledge input. Specifically, five If-Then rules were used whose IF components (aka rule firing strengths) were derived from the expert knowledge from the decision tree in 1) abovementioned. The consequent (i.e., Else) parameters of ANFIS were optimized via backpropagation [30]. The reader can refer to [33] [34], and [35] for an in-depth discussion of the mathematics, optimization, and robust possibilistic clustering-based initialization of ANFIS. Finally, the output decision threshold was chosen through DTA.

The different *LPs* detector types were tested independently and in combination during the fusion step with the other three component types (i.e., *missiles*, *TELs*, and *TEL Groups*).

1) Empty *LPs* plus three (*Empty LPs+3*).
2) Combined *LPs* plus three (*Combo LPs+3*).
3) *Empty LPs* and *combo LPs* plus three (All 5).

C. MLP Input Data Normalization

We found that the MLPs had some difficulty training with datasets that had larger values, so we used the common practice of linearly scaling and bounding to constrain the data to fall within the range $[-1, 1]$. Let v_i be the vector of values over the entire dataset for component i for a given feature and let t_i be the DTA thresholds computed for component i, then the normalized and bounded vector v_i' can be defined as follows:

$$v_i' = \begin{cases} \text{if} & (v_i - t_i)/t_i \in [-1, 1], \text{ then } (v_i - t_i)/t_i \\ \text{if} & (v_i - t_i)/t_i < -1, \text{ then } -1 \\ \text{if} & (v_i - t_i)/t_i > 1, \text{ then } 1. \end{cases} \quad (4)$$

D. Results & Observations

Over 200 different combinations of data feature types, component combinations, and fusion techniques were tested in this study to improve the final detection of candidate *SAM Sites*.

Evaluation of the *F1* score improvements (Table VI) shows that decision-level component fusion can reduce the relative error rate by up to 96.75%. It was somewhat surprising that the raw count feature generated five out of the top six best results. Although *Combo LPs* were only able to generate an *F1* score of 68.4% using DTA, the neural approaches (MLP and ANFIS) were able to do slightly better using multiple components where the top results fused all five components in an MLP to yield an *F1* score of 71.4%. Comparisons of *F1* scores for different feature types and fusion techniques can be found in Fig. 6.

However, when performing a broad area search for a very rare object (low geographic occurrence rate), it is often desirable to sacrifice some error reduction in order to achieve a higher *TPR*. The results in Table VII show that the highest *F1* score is 45.1% while achieving a *TPR* of 100%. Although this *F1* score is less than half of the maximum in Table VI, this technique still achieved a 88.5% relative error reduction compared to the baseline (no component fusion) results for the candidate *SAM Site* locations within the SE China AOI. These scores were produced using custer count features and the all five component combination as inputs to a simple MLP. It is also worth noting that four of the top five scores used the *empty LPs+3* component combination. Comparisons of TPRs for different feature types and fusion techniques can be found in Fig. 7.

It was also observed that cluster score truncation and normalization was able to improve the *F1* scores for DTA when fusing multiple component detectors. However, the introduction of negative score penalty did not improve the score further (Fig 8), while introducing expert weighting (described in Section VI-A) also showed no improvement for the *F1* scores.

TABLE VI

EXPERIMENT RESULTS WITH HIGHEST *F1* SCORES. THE FIRST LINE AFTER THE HEADER (IN RED) ARE THE RESULTS FOR *SAM SITE* DETECTION WITHOUT ERROR REDUCTION FROM SPATIAL FUSION OF ANY COMPONENT FEATURE TYPE(S). THE HIGHEST *F1* SCORES WERE ACHIEVED BY FUSING MULTIPLE COMPONENTS USING NEURAL LEARNING TECHNIQUES (MLP OR ANFIS). ALSO, RAW DETECTION COUNTS (PRECLUSTERING) SHOWED THE MOST SEPARABILITY. ALL TOP SOLUTIONS ACHIEVED A RELATIVE ERROR REDUCTION OF GREATER THAN 96%. THESE RESULTS WOULD BE OPTIMAL IF ERROR REDUCTION WAS THE PRIMARY GOAL. THE ERROR RATE INCLUDES BOTH FALSE POSITIVES AND FALSE NEGATIVES

Components	Feature Type	Processing Technique	Component Negatives	*TP*	*FP*	*TPR* (Recall)	*PPV* (Precision)	*F1* score	Error / km^2 (x10^{-3})	Relative Error Reduction
SAM Sites	BASELINE-NO COMPONENTS			16	338	100.00%	4.52%	8.65%	3.080	n/a
All 5	Raw Counts	MLP	NO	15	11	93.75%	57.69%	71.43%	0.109	96.45%
Combo LPs+3	Raw Counts	ANFIS	NO	13	8	81.25%	61.90%	70.27%	0.100	96.75%
Combo LPs+3	Raw Counts	MLP	NO	14	10	87.50%	58.33%	70.00%	0.109	96.45%
Combo LPs	Cluster Count	DTA	n/a	13	9	81.25%	59.09%	68.42%	0.109	96.45%
Combo LPs	Raw Count	DTA	n/a	13	9	81.25%	59.09%	68.42%	0.109	96.45%
All 5	Raw Count	ANFIS	NO	13	9	81.25%	59.09%	68.42%	0.109	96.45%

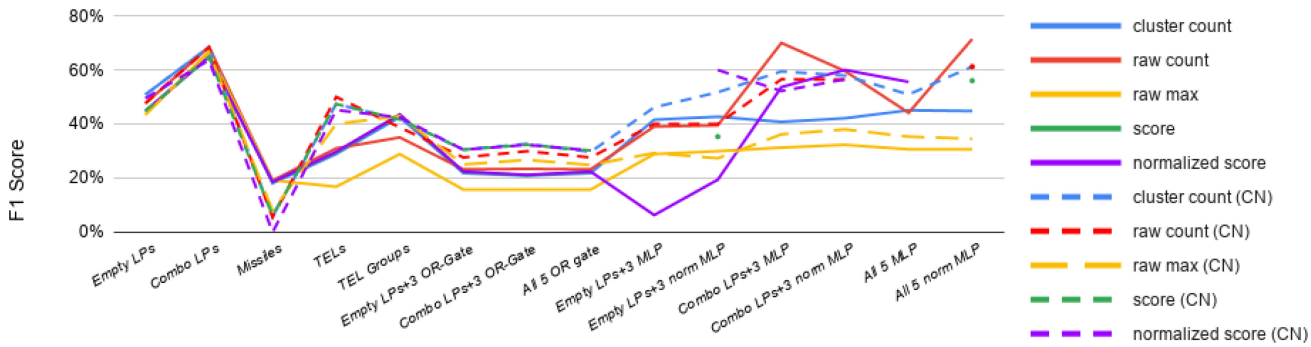

Fig. 6. Comparison of *F1* scores produced for candidate *SAM Site* locations from different fusion techniques. Techniques include individual component threshold from DTA as well as component fusion using an OR gate and MLP. Note that "(CN)" at the end of the feature type label in the key indicates that component negative models were used in the processing. Gaps in scores occur when the MLP was unable to train on a given feature type.

TABLE VII

EXPERIMENT RESULTS WITH HIGHEST *F1* SCORES WHILE MAINTAINING A *TPR* OF 100%. THE FIRST LINE AFTER THE HEADER (IN RED) ARE THE RESULTS FOR *SAM SITE* DETECTION WITHOUT ERROR REDUCTION FROM SPATIAL FUSION OF ANY COMPONENT FEATURE TYPE(S). THE HIGHEST *F1* SCORES RESULTED FROM FUSING A FEATURE FROM ALL COMPONENTS WITH A SIMPLE MLP. ALSO, CLUSTER COUNT FEATURES YIELDED THE TOP RESULTS. ALL TOP SOLUTIONS SHOW A REDUCTION OF RELATIVE ERROR BETWEEN $85.2\% - 88.5\%$ WHICH IS $3X$ THE ERROR RATE SHOWN IN TABLE VII

Components	Feature Type	Processing Technique	Component Negatives	*TP*	*FP*	*TPR* (Recall)	*PPV* (Precision)	*F1* score	Error/ km^2 (x10^{-3})	Relative Error Reduction
SAM Sites	BASELINE-NO COMPONENTS			16	338	100%	4.52%	8.65%	3.080	n/a
All 5	Cluster Count	MLP	NO	16	39	100%	29.09%	45.07%	0.355	88.46%
Empty LPs+3	Cluster Count	MLP (Normalized)	NO	16	43	100%	27.12%	42.67%	0.392	87.28%
Empty LPs+3	Cluster Count	MLP	NO	16	45	100%	26.23%	41.56%	0.410	86.69%
Empty LPs+3	Raw Count	MLP (Normalized)	YES	16	48	100%	25.00%	40.00%	0.437	85.80%
Empty LPs+3	Raw Count	MLP	NO	16	50	100%	24.24%	39.02%	0.456	85.21%

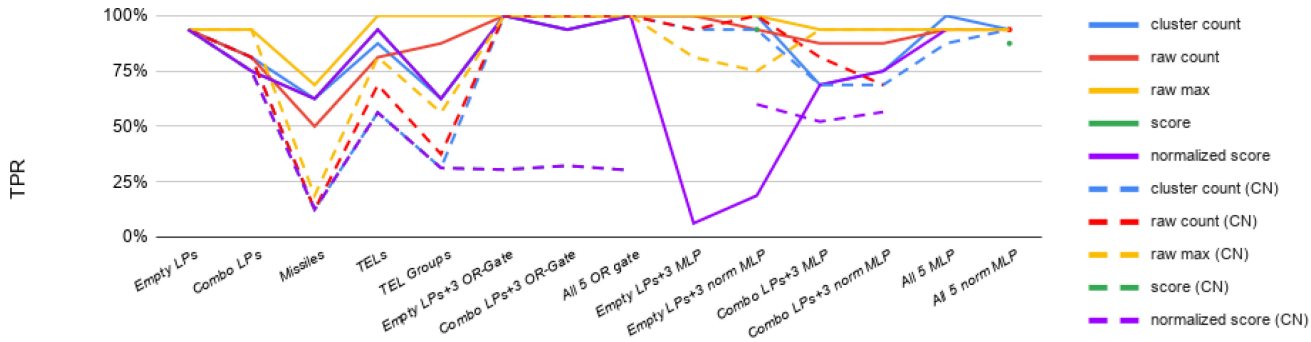

Fig. 7. Comparison of TPR produced for candidate *SAM Site* features from different fusion techniques.

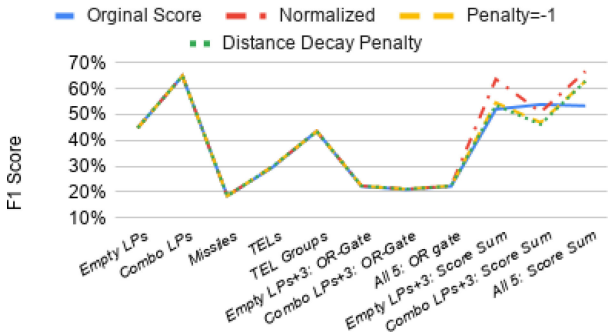

Fig. 8. *F1* score results for DTA thresholds of original cluster scores, normalized cluster scores, cluster scores with a penalty of –1 and distance-decay penalty with $R = 150$ m.

TABLE VIII
AVERAGE RANK OF KNOWN *SAM SITES* (TPS) IN SE CHINA AOI FROM FUSING CLUSTER SCORES FROM A SINGLE COMPONENT OBJECT CLASS WITH A BASELINE CANDIDATE *SAM SITE* CLUSTER SCORE

		with Single Component Fusion				
	SAM Site Only	*Empty LPs*	*Combo LPs*	*Missiles*	*TELs*	*TEL Groups*
ResNet-101 [19]	139.9	n/a	n/a	n/a	n/a	n/a
NASNet	62.8	36.4	40.8	43.0	28.0	46.1
w/ Norm	62.8	34.3	34.4	43.6	28.1	47.3
w/ Norm & Penalty	62.8	34.0	34.3	43.6	**27.9**	47.1
Including Component Negatives						
w/ Norm	n/a	n/a	n/a	79.1	28.8	51.6
w/ Norm & Penalty	n/a	n/a	n/a	79.1	28.7	51.48

Additionally, in general, there was improvement in *F1* scores for models trained with component negatives, however these improvements came at a sacrifice in *TPR* and only have one appearance in Tables VI and VII. This can be interpreted as ambiguity being introduced to the dataset by essentially asking the detector to ignore the background (i.e., the *LP*) and focus on the smaller component.

VI. COMPONENT METRICS FUSION FOR IMPROVING CANDIDATE *SAM SITE* RANKINGS

This section discusses techniques, observations, and results used to rerank candidate *SAM Sites* for utilization in machine-assisted human analytic workflows. The objective is to utilize the component detection clusters to rerank the candidate *SAM Sites* such that true *SAM Sites* appear higher in a rank-ordered list relative to a baseline ranking derived only from the candidate *SAM Sites'* cluster scores (Table VIII). An overview of the processing flow is given in Fig. 9.

A. Candidate Site and Component Score Spatial Fusion

Normalized cluster scores for candidate *SAM Sites* and all components found within *R* are summed using uniform or human expert provided weights (Fig. 9). Expert weights were only used when fusing all four components with its corresponding candidate *SAM Site*. The weights were: four for *LP*, two for *TEL Groups*, and one for *missiles*, *TELs*, and *SAM Sites*.

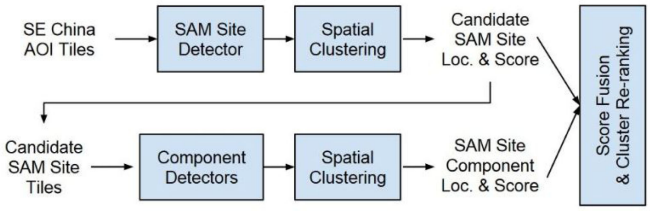

Fig. 9. Process flow used for improved ranking of candidate *SAM Sites*.

TABLE IX
AVERAGE RANK OF KNOWN *SAM SITES* (TPS) IN SE CHINA AOI FROM FUSING CLUSTER SCORES FROM ALL FOUR COMPONENT OBJECT CLASSES WITH THE BASELINE CANDIDATE *SAM SITE* CLUSTER SCORE

	Empty LPs		*Combo LPs*	
	Unweighted Fusion	Weighted Fusion	Unweighted Fusion	Weighted Fusion
NASNet	26.3	21.4	25.3	22.9
w/ Norm	20.3	22.9	17.9	**15.9**
w/ Norm & Penalty	19.9	22.5	17.8	16.0
Including Component Negatives				
w/ Norm	24.8	24.9	18.1	16.8
w/ Norm & Penalty	24.1	24.9	18.1	16.8

B. Results and Observations

The *TEL* detector rendered the most improvement in the average cluster rank of known *SAM Sites* (*TPs*) compared to fusion with any other single component detector (Table VIII). This, coupled with the *Combo LPs* detector and other component detectors trained with expert weighting (Section VI-A) improved the average cluster rank of known *SAM Sites* (*TPs*) to 15.9 (Table IX). This is $\sim 4X$ better than the average rank for *SAM Sites* without spatial fusion of the component object cluster scores.

We observed that the addition of normalization and penalty had no detectable impact on the known *SAM Site* TP average cluster rank. This indicates minimal FP presence and/or uniformly distributed FP noise within the candidate *SAM Site* locations generated by the spatial clustering algorithm.

Component negative models improved the ranking results compared to the *SAM Site* score alone, but not as well as models trained without component negatives. Again, this can be interpreted as ambiguity being introduced to the dataset by essentially asking the detector to ignore the background (i.e., the *LP*) and focus on the smaller component.

VII. CONCLUSION AND FUTURE WORK

This study extended the work in [19] where a combination of a DNN scanning and spatial clustering was used to perform a machine-assisted broad area search and detection of *SAM Sites* in a SE China AOI of ~90 000 km^2.

Here, we significantly improved upon this prior study by using multiple DNNs to detect smaller component objects, e.g., *LPs*, *TELs*, etc., belonging to the larger and more complex *SAM Site* feature. Scores computed from an enhanced spatial clustering algorithm were normalized to a reference space so

that they were independent of image resolution and DNN input chip size. A variety of techniques were then explored to fuse the DNN detections from the multiple component objects to improve the final detection and retrieval (ranking) of DNN detections of candidate *SAM Sites*. Key results from this effort include.

1) Spatial fusion of DNN detections from multiple component objects using neural learning techniques that maximize the *F1* score reduced an initial set of ∼350 *SAM Site* detections (Table V) to only ∼25 candidate *SAM Sites* (Table VI).
2) An alternate spatial fusion approach from that used in 1) reduced the overall error rate by >85% while preserving a 100% *TPR* (Table VII) and also reduced the initial set of detections to ∼55–60 candidate *SAM Sites*.
3) The average rank of 16 known *SAM Sites* (*TP*s) in a list of ∼350 candidate *SAM Sites* was improved by $\sim 9X$ (Tables VIII and IX) compared to the previous study [19].

In future work, we plan to apply this approach to a variety of other challenging object search and detection problems in large-scale remote sensing image datasets, investigate data-driven optimization of the component fusion weights and compare performance versus human-expert provided weights, extend this approach to include fusion of multitemporal DNN detections, extend this approach to include fusion of multisource DNN detectors applied to high-resolution electro-optical (EO)/mulitspectral (MS) and synthetic-aperture radar (SAR) imagery, and explore how to use more sophisticated fusion techniques (similar to ANFIS) to maintain *TPR* while achieving even higher error reduction.

REFERENCES

[1] G. J. Scott, K. C. Hagan, R. A. Marcum, J. A. Hurt, D. T. Anderson, and C. H. Davis, "Enhanced fusion of deep neural networks for classification of benchmark high-resolution image datasets," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 9, pp. 1451–1455, Sep. 2018.

[2] J. A. Hurt, G. J. Scott, D. T. Anderson, and C. H. Davis, "Benchmark meta-dataset of high-resolution remote sensing imagery for training robust deep learning models in machine-assisted visual analytics," in *Proc. IEEE Appl. Imagery Pattern Recognit. Workshop*, Oct. 2018, pp. 1–9.

[3] G. J. Scott, M. R. England, W. A. Starms, R. A. Marcum, and C. H. Davis, "Training deep convolutional neural networks for land cover classification of high-resolution imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 549–553, Apr. 2017.

[4] G. J. Scott , R. A. Marcum, C. H. Davis, and T. W. Nivin, "Fusion of deep convolutional neural networks for land cover classification of high-resolution imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 9, pp. 1638–1642, Sep. 2017.

[5] Y. Yang and S. Newsam, "Bag-of-visual words and spatial extensions for land-use classification," in *Proc. ACM SIGSPATIAL Int. Conf. Adv. Geogr. Inf. Syst.*, 2010, pp. 270–279.

[6] G. Sheng, W. Yang, T. Xu, and H. Sun, "High-resolution satellite scene classification using a sparse coding based multiple feature combination," *Int. J. Remote Sens.*, vol. 33, no. 8, pp. 2395–2412, 2012.

[7] G. Cheng, J. Han, and X. Lu, "Remote sensing image scene classification: Benchmark and state-of-the-art," *Proc. IEEE*, vol. 105, no. 10, pp. 1865–1883, Oct. 2017.

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 779–788.

[9] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 580–587.

[10] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1440–1448.

[11] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[12] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," 2018, *arXiv:1804.02767*.

[13] W. Liu *et al.*, "SSD: Single shot multibox detector," in *Proc. 14th Eur. Conf. Comput. Vis.*, 2016, vol. 9905, pp. 21–37.

[14] J. Shermeyer and A. V. Etten,"The effects of super-resolution on object detection performance in satellite imagery," to be published, 2019.

[15] Y. Koga, H. Miyazaki, and R. Shibasaki, "A method for vehicle detection in high-resolution satellite images that uses a region-based object detector and unsupervised domain adaptation," *Remote Sens.*, vol. 12, 2020, Art. no. 1068.

[16] Z. Xin, H. Liangxiu, H. Lianghao, and Z. Liang, "How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery?" *Remote Sens.*, vol. 12, 2020, Art. no. 417.

[17] S. P. DelMarco, V. Tom, H. Webb, W. Snyder, C. Jarvis, and D. Fay, "Shape-based ATR for wide-area processing of satellite imagery," *Proc. SPIE*, vol. 10988, 2019, Art. no. 1098805.

[18] Y. Yanan, L. Zezhong, R. Bohao, C. Jingyi, L. Sudi, and L. Fang, "Broad area target search system for ship detection via deep convolutional neural network," *Remote Sens.*, vol. 11, 2019, Art. no. 1965.

[19] R. A. Marcum, C. H. Davis, G. J. Scott, and T. W. Nivin, "Rapid broad area search and detection of chinese surface-to-air missile sites using deep convolutional neural networks," *J. Appl. Remote Sens.*, vol. 11, no. 4, 2017, Art. no. 042614.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[21] B. Zoph, V. Vasudevan, J. Shlens, and Q. Le, "Learning transferable architectures for scalable image recognition," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 8697–8710.

[22] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1800–1807.

[23] H. Cai, L. Zhu, and S. Han, "ProxylessNAS: Direct neural architecture search on target task and hardware," in *Proc. Int. Conf. Learn. Representations*, 2019, pp. 1–13. [Online]. Available:https://openreview.net/forum?id=HylVB3AqYm

[24] M.Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 6105–6114.

[25] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 248–255.

[26] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," Dec. 2014, *arXiv:1412.6980*.

[27] A. B. Cannaday II, R. L. Chastain, J. A. Hurt, C. H. Davis, G. J. Scott, and A. J. Maltenfort, "Decision-level fusion of DNN outputs for improving feature detection performance on large-scale remote sensing image datasets," in *Proc. IEEE Int. Conf. Big Data* , Los Angeles, CA, USA, 2019, pp. 5428–5436.

[28] N. Ye, K. M. A. Chai, W. S. Lee, and H. L Chieu, "Optimizing F-measures: A tale of two approaches," in *Proc. Int. Conf. Mach. Learn.*, 2012, pp. 289–296.

[29] D. D. Lewis, "Evaluating and optimizing autonomous text classification systems," in *Proc. 18th Annu. Int. ACM SIGIR Conf. Res. Develop.*, 1995, pp. 246–254.

[30] J. Jang, "ANFIS adaptive-network-based fuzzy inference system," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, no. 3, pp. 665–685, May/Jun. 1993.

[31] A. Abraham, "Adaptation of fuzzy inference system using neural learning," in *Fuzzy Systems Engineering: Theory and Practice, Studies in Fuzziness and Soft Computing*, N. Nedjah, L. de Macedo Mourelle Eds. Berlin, Germany: Springer, 2005, pp. 53–83.

[32] D. Karaboga and E. Kaya, "Adaptive network based fuzzy inference system (ANFIS) training approaches: A comprehensive survey," *Artif. Intell. Rev.*, vol. 52, pp. 2263–2293, Jan. 2018.

[33] B. Ruprecht *et al.*, "Fuzzy logic-based fusion of deep learners in remote sensing," in *Proc. FuzzIEEE*, 2019.

[34] B. Ruprecht *et al.*, "Neuro-fuzzy logic for parts-based reasoning about complex scenes in remotely sensed data" *Signal Process., Sensor/Informat. Fusion, Target Recognit. XXIX*, pp. 61–63, 2020.

[35] B. Ruprecht *et al.*, "Possibilistic clustering enabled neuro fuzzy logic," in *Proc. World Congr. Comput. Intell.*, 2020.

Alan B. Cannaday II (Student Member, IEEE) received the B.S. degree in mathematics and the M.S. degree in computational engineering and science from the University of Utah, Salt Lake City, UT, USA, in 2010 and 2012, respectively. He is currently working toward the Ph.D. degree in computer science with the University of Missouri, Columbia, MO, USA.

He worked for five years as a Research Engineer for FamilySearch where his work focused on historical document analysis and automatic content extraction. Since entering the Ph.D. program in 2018, he has worked as a Graduate Research Assistant for the Center for Geospatial Intelligence with studies and papers focused in computer vision, neural networks, object detection, fuzzy data fusion, and remote sensing.

Curt H. Davis (Fellow, IEEE) was born in Kansas City, MO, on October 16, 1964. He received the B.S. and Ph.D. degrees in electrical engineering from the University of Kansas, Lawrence, KS, USA, in 1988 and 1992, respectively.

He is currently the Naka Endowed Professor of Electrical Engineering and Computer Science with the University of Missouri, Columbia, MO, USA and Director of the Center for Geospatial Intelligence which he founded in 2005. Over his 30-year scientific career his research has involved the use of satellite microwave and optical remote sensing systems for numerous applications in the areas of earth observation and science including ice sheet monitoring, change detection, geospatial information processing, and national security. His current research is focused on the development and application of deep machine learning methods for object detection and scene recognition to improve analytic workflows involving the search, discovery, retrieval, and/or monitoring of objects and features of interest in large volumes of high-resolution satellite imagery. His research results have been documented more than 60 refereed journal publications and 90 symposia presentations and proceedings. His most significant scientific studies have been published in top scientific journals like *Science*, *Nature*, and *the Journal of Geophysical Research*, while the majority of his technical contributions to the field of remote sensing have been published in the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Dr. Davis was named an IEEE Fellow for his "contributions to satellite remote sensing", in 2008 and has was the recipient of numerous other awards throughout his career. Examples include the NSF Antarctica Service Medal (1988 and 1989), International Union of Radio Science Young Scientist Award (1996), and the NASA New Investigator Program (1996–1999). He has been serving as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, since 2005.

Grant J. Scott (Senior Member, IEEE) received the B.S. and M.S. degrees in computer science and the Ph.D. degree in computer engineering and computer science from the University of Missouri, Columbia, MO, USA, in 2001, 2003, and 2008, respectively.

He is a founding Director of the Data Science and Analytics Masters Degree program with the University of Missouri. He is an Assistant Professor with the Electrical Engineering and Computer Science Department, University of Missouri. He is exploring novel integrations of computational hardware and software to facilitate high performance advances in large-scale data science, computer vision, and pattern recognition. He has participated in a variety of professional networking and academic events, as well as worked with a variety of groups to bring data science training to their people (MU Public Policy, USDA, IEEE international conferences). His current research efforts encompass areas such as: real-time processing of large-scale sensor networks, parallel/distributed systems, and Internet of Things data; deep learning technologies applied to geospatial data sets for land cover classification and object recognition; extensions of enterprise relational database management system (RDBMS) with high preformance computing (HPC) coprocessors; crowd-source information mining and multimodal analytics; high performance and scalable content-based retrieval (geospatial data, imagery, biomedical); imagery and geospatial data analysis, feature extraction, object-based analysis, and exploitation; pattern recognition databases and knowledge-driven high-dimensional indexing; and image geolocation. He has leveraged this experience in the development of innovative remote sensing (satellite and airborne) change detection technologies, resulting in five US Patents.

Blake Ruprecht (Student Member, IEEE) is currently working toward the master's degree in computer science with the Electrical Engineering and Computer Science Department, University of Missouri, Columbia, MO, USA.

His research interests include explainable artificial intelligence, computational intelligence, neuro-fuzzy systems, AI risk research, and effective altruism.

Derek T. Anderson (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Missouri, Columbia, MO, USA, in 2010.

He is an Associate Professor in electrical engineering and computer science with the University of Missouri and directs the Mizzou Information and Data Fusion Laboratory. He has authored more than 150 articles. His research is information fusion in computational intelligence for signal/image processing, computer vision, and geospatial applications.

Dr. Anderson was the recipient of the Best Overall Paper Award at the 2012 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), and the 2008 FUZZ-IEEE Best Student Paper Award. He was the Program Co-Chair of FUZZ-IEEE 2019, an Associate Editor for the IEEE TRANSACTIONS ON FUZZY SYSTEMS, Vice Chair of the IEEE CIS Fuzzy Systems Technical Committee, and an Area Editor for the *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*.