Fuzzy aggregation for multimodal remote sensing classification

Kristen Nock
U.S. Naval Research Laboratory
Washington, DC, USA
kristen.nock@nrl.navy.mil

Elizabeth Gilmour
U.S. Naval Research Laboratory
Washington, DC, USA
elizabeth.gilmour@nrl.navy.mil

***Abstract*—This paper investigates methods of fusing hyperspectral imagery (HSI) and LiDAR data for urban land use and land cover classification. A variety of fusion methods including combination rules, deep neural networks, and fuzzy aggregation are compared against using any single modality for classification. The experimental results demonstrate that the two fuzzy aggregation methods, the linear order statistic neuron (LOSN) and the Choquet integral (ChI), achieved the best overall and average classification accuracy, respectively. We further discuss how the fuzzy aggregation methods provides advantages with difficult samples and the opportunity to gain network explainability.**

***Index Terms*—hyperspectral imagery, LiDAR, Choquet integral, fuzzy aggregation, remote sensing**

I. Introduction

Remote sensing data collected by airplanes flying over the surface of the earth has uses ranging from agriculture to urban planning to disaster response, but the interpretation of remote sensing images is necessary to extract useful information. Remote sensing data provide rich information about land use and other natural phenomena through various sensor types. Hyperspectral imagery (HSI) and LiDAR are two remote sensing data types that will be examined in this paper. Each of the data types provides vital but incomplete information for land use classification. HSI provides information about material characteristics, so it can distinguish between parking lots and fields; however, it cannot be used to differentiate between objects where the same materials occur at different heights [1]. LiDAR, in contrast, can detect surface properties like height and roughness. For example, LiDAR cannot be used to differentiate between objects with the same elevation and surface roughness that are made of different materials [1], like artificial and natural grass, but can easily distinguish between the height of a parking garage and the height of a parking lot.

Intuitively, the fusion of information from both HSI and LiDAR would combine complementary information from the two sensor modalities to make a more informed decision [2]. This sort of sensor fusion can be thought of as consulting experts with different specialties and considering their opinions. However, developing frameworks to combine the sensor inputs and decisions remains a challenge.

The authors would like to thank the National Center for Airborne Laser Mapping and the Hyperspectral Image Analysis Laboratory at the University of Houston for acquiring and providing the data used in this study, and the IEEE GRSS Image Analysis and Data Fusion Technical Committee. This research was sponsored by the Naval Research Laboratory's Base Program, and the Office of Naval Research. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. Government.

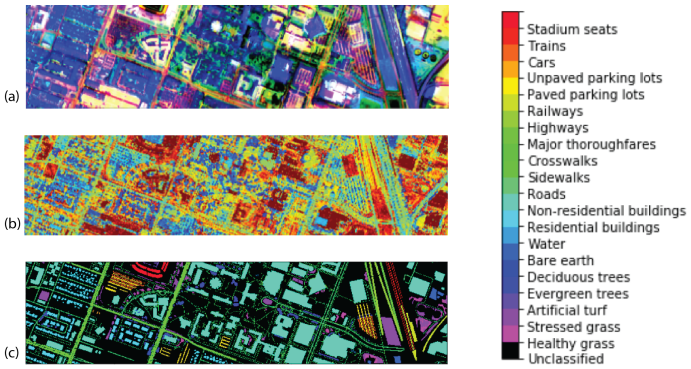

Fig. 1. (A) The hyperspectral data (HSI) covering a 380-1050 nm spectral range with 48 bands at a 1m ground sampling distance. (B) The multispectral LiDAR point cloud data, composed of intensity raster and digital surface models with a resolution of 0.5m ground sampling distance. (C) Corresponding ground truth data containing 20 urban land use and land cover classes.

Previously, fuzzy aggregation has been used for fusing predictions from different network architectures in ensemble models in remote sensing studies [3]. However, this study is the first application of fuzzy aggregation, in the form of the linear order statistic neuron (LOSN) and the Choquet integral (ChI), to perform multimodal sensor fusion of HSI and LiDAR data. Further, this study compares the accuracy of these fuzzy aggregation methods to the use of combination rules and fusion within a deep neural network.

II. Background

LiDAR is a type of data recorded from pulsed lasers reflected off the earth's surface. The returning light is collected, and the phase changes, wavelengths, and angles of the returning light are used to create an image. LiDAR provides information about terrain and vegetation that cannot be obtained from optical sensors. LiDAR has important uses in topography, such as flood plain analysis, and it is also used in urban planning as it provides detailed information about structures and building footprints.

Hyperspectral remote sensing uses spectroscopy to obtain material characteristics. Objects absorb light from the sun and then emit light at different frequencies, depending on their material characteristics. The hyperspectral imaging camera acquires the emitted light in frequencies ranging from the infrared to the visible light, producing a type of imagery with the capability to distinguish physical materials. HSI has a variety of uses, including use in precision agriculture to study crop health, in ecology to study land use changes, and in military applications for target detection [4].

HSI produces a data cube with spatial dimensions related to the area covered and a depth that represents the many spectral channels collected. The data is difficult to visually interpret, so it is typically classified through segmentation in which the class of each pixel is predicted. The classes represent different materials or land use types. Numerous types of supervised classification models have been used on HSI, including simple look-up tables, decision trees [5], random forests [6], and support vector machines [7]. In recent years, deep learning frameworks have become important for classification because of their powerful feature extraction [8].

Deep learning classification of HSI has typically focused either on the spectral features or the spatial features. Spatial feature networks, which use deep neural networks to extract spatial features of HSI, can provide good classification accuracy even though they ignore the important spectral features [8]. Spectral feature networks are neural networks that use the many spectral bands of HSI. These spectral characteristics are arguably the most important characteristic of HSI, and can be used to perform pixel-wise classification. Each pixel of the HSI is represented as a spectral vector which contain the number of spectral bands, or the depth of the hyperspectral image cube. Due to the challenging and redundant information in the spectral data, spectral feature networks do not necessarily maximize classification accuracy [8]. Additionally, as more bands are added to the HSI, the classes become more separable but more training data is needed to accurately classify them [4]. The Geoscience and Remote Sensing Society (GRSS) sponsors data fusion contests, challenging teams to develop fusion frameworks for HSI and LiDAR. In a 2013 data fusion contest, GRSS provided co-registered HSI and LiDAR that was collected over the University of Houston campus [9]. Winning teams developed frameworks that included supervised and unsupervised methods as well as handcrafted feature extraction. While the end result is automated land use classification, the process involves custom model techniques that do not generalize well.

In 2018, the GRSS data fusion contest was repeated with an updated dataset collected over the University of Houston. This multimodal dataset of urban land use acquired from an area of downtown Houston was released as part of the 2018 GRSS Data Fusion Contest [10]. The dataset, grss_dfc_2018 [11], includes co-registered LiDAR and HSI, as well as high-resolution imagery (Fig. 1). The hyperspectral data includes 48 bands covering a spectral range of 380-1050 nm with a 1m ground sampling distance [10]. The LiDAR data is composed of intensity raster and digital surface models with 0.5m ground sampling distance. The dataset has 20 classes related to urban objects and land use. In the contest, the majority of the scene was used for training, and a portion was held aside for testing.

In contrast to the sparse ground truth dataset used for 2013 contest, the 2018 data was dense to promote the advancement of deep learning-based approaches. In 2018, deep learning-based approaches overwhelmingly occupied the leaderboard. Most of the top-ranked teams used several different neural networks trained together. However, the best performing approaches were based on deep neural networks along with post processing and object detection [10]. Post-processing, including hard-coded rather than learned methods, was needed to correct systematic errors in the deep learning classifications.

While fuzzy multiple classifier systems for HSI and LiDAR data have been used in the past [12], we further hypothesize that fuzzy set theory in the form of the LOSN and the ChI can reduce the dependence on hand-crafted post-processing in multimodal sensor fusion of HSI and LiDAR. Instead of a one size fits all operator (e.g., average, max, etc.), data-driven fuzzy aggregation can intelligently learn contextual combination logic, deciding a unique and appropriate operator for each of the classes present. This study builds off of deep neural networks developed by others, but combines them in a novel framework and tests different methods for aggregating their decisions. It should be noted that this study did not have access to the 2018 GRSS contest test data, so the results cannot be directly compared.

III. Technical Approach

To classify the co-registered HSI and LiDAR, we test different types of neural networks and different aggregation methods to combine the decisions of the neural networks. These methods include single sensor networks, unified networks, combination rules, and fuzzy aggregation. All methods we tested are built on three neural networks, the convolutional neural network (CNN), multi-scale convolutional neural network (MSCNN), and long short-term memory (LSTM). The spatial networks, the CNN and MSCNN, are trained with patches of nine pixel by nine pixel imagery where the central pixel determines the class label. In contrast, the LSTM is trained using all spectral bands for a single pixel. The three neural networks and the fusion framework are shown in Fig 2. For all models, a methodical manual search was performed over ranges of hyperparameters to optimize model performance.

A. Convolutional neural network

An important deep learning model, the convolutional neural network (CNN), is inspired by the structure of the human visual system. The structure of the CNN consists of a stack of alternating convolutional layers and pooling layers followed by fully connected (FC) layers. This structure creates a convolutional feature map and makes use of location connections to extract contextual 2-D spatial features. In the convolution layers, images are convolved with a set of learned filters.

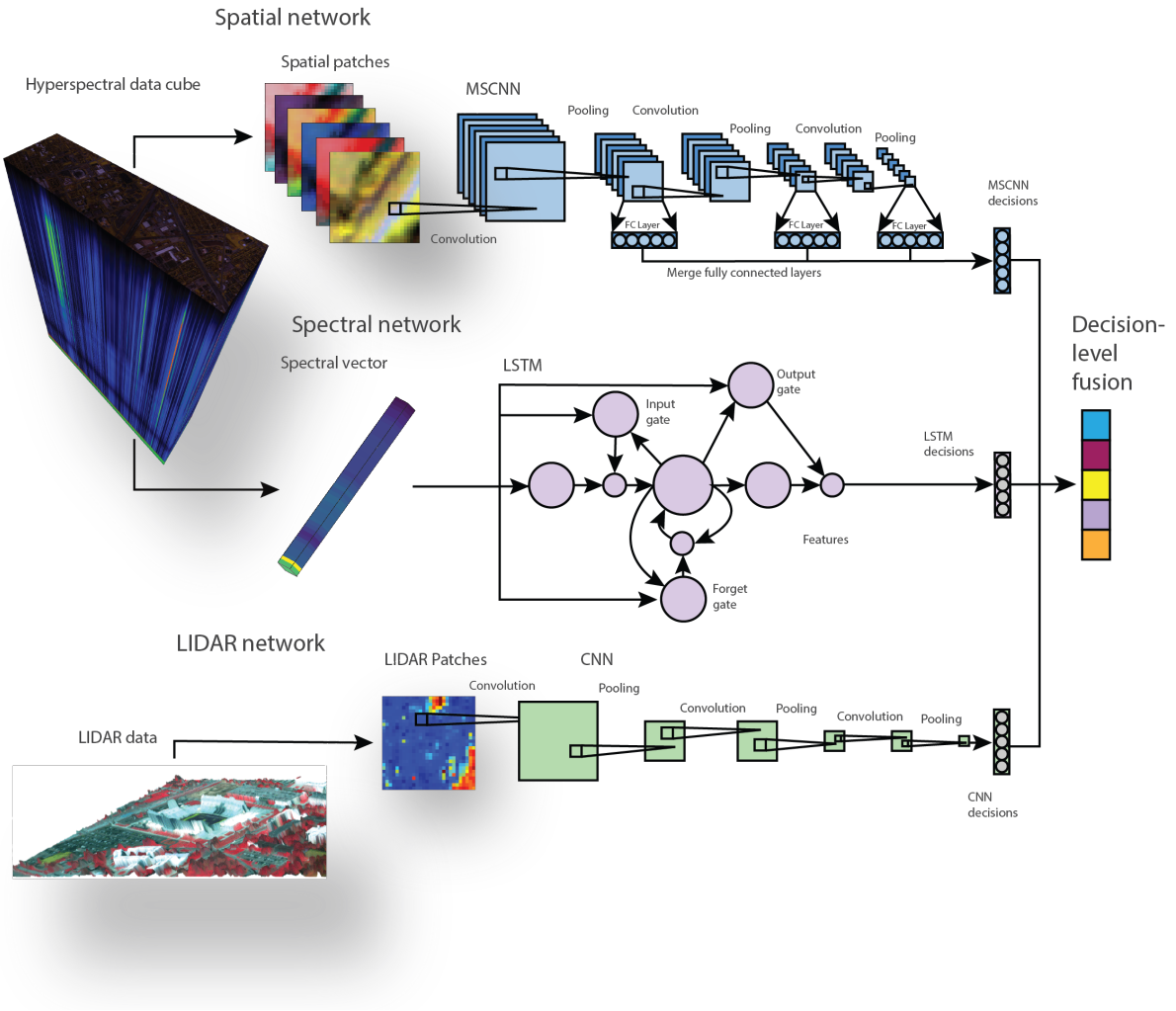

Fig. 2. Three separate neural network feeds are fused to make predictions. The spatial feature network (top) is multi-scale convolutional neural network (MSCNN) that is trained on nine pixel by nine pixel patches of hyperspectral imagery (HSI). The spectral feature network (middle) is an LSTM that is trained on single pixel vectors of HSI. The LiDAR network (bottom) is a CNN that is trained on nine pixel by nine pixel patches of LiDAR. These three neural networks are combined using different fusion methods to predict the class of training samples.

Pooling layers are used to reduce redundant information, and the FC layers are used to flatten the feature maps into n-dimensional vectors to predict the labels. We split the data into patches to perform patch-based classification. Nine pixel by nine pixel patches of LiDAR were classified with a CNN.

B. Spatial feature network

In traditional CNNs, the final FC layers are used to make the final predictions. In contrast, the multi-scale convolutional neural network (MSCNN), introduced by Xu et al. in 2018 [13], learns hierarchical spatial features from patches of HSI data at many layers. The MSCNN is trained with HSI patches where the central pixel determines the class label. The FC layers after each convolutional layer are concatenated, and this combined layer is used for the final feature vector. As a result, this network uses features from multiple spatial scales to make predictions. This improves upon many other spatial feature neural networks for HSI classification as it can use several different spatial scales to predict the class.

C. Spectral feature network

Developed for time series predictions, long short-term memory (LSTM) is a type of recurrent neural network that has also shown success classifying hyperspectral imagery [13]. Each pixel of the HSI is represented as a vector of length N where N is the number of spectral bands. The spectral vectors have a sequence-like data structure, so even though the spectral vectors do not contain time series data, the LSTM model can also be used to analyze this data type [13]–[15].

D. Fusion methods

This paper proposes a new fusion architecture where three separate neural networks are combined using decision level fusion to classify the data. The decisions of the MSCNN

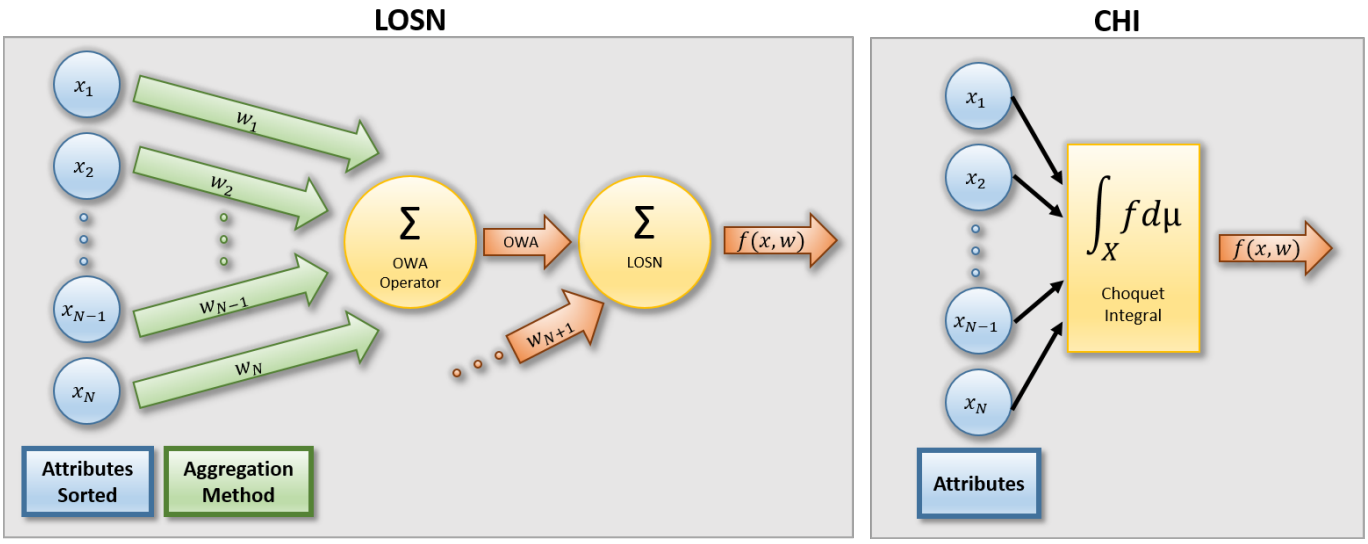

Fig. 3. This diagram shows the fuzzy aggregation operators used. The linear order statistic neuron (LOSN) is a generalization of the ordered weighted average (OWA) operator. Mathematically, the LOSN is the sum of a bias, and dot product of a sorted input and weights selected based on a aggregation method. The Choquet integral (ChI) acts as a generalized expectation operator. There will be N number of non-monotonic fuzzy measures μ used in this parametric non-linear aggregation function.

as the HSI spatial feature network, the LSTM as the HSI spectral feature network, and a CNN for the LiDAR are fused. The two data types each have strengths and weaknesses for a classification problem, so combining them efficiently can increase classification accuracy. We test three different methods to find the best way to combine the classifiers and perform data fusion.

a) Combination rules: Combination rules are simple algebraic rules for deciding which prediction to trust. They are based on the predictions of the classifiers and the confidence of each classifier's prediction. Simple voting looks at the predictions from each of the classifiers, and chooses the prediction that the majority of classifiers picked [2], [3]. Let $d_{(t,j)}(x) \in [0,1]$ be the decision of classifier t for class j relative to data point/input x. The chosen class (w^*) is

$$w^* = mode(\underset{j=1,\ldots,C}{\operatorname{argmax}}\, d_{(1,j)}(x), ..., \underset{j=1,\ldots,C}{\operatorname{argmax}}\, d_{(T,j)}(x)) \quad (1)$$

where T is the number of classifiers and C is the number of classes. In other words, the mode (the class prediction that appears most often) is used.

Maximum fusion is based on the confidence of the predictions. The prediction with the highest confidence among all of the networks is chosen [2], [3]. Let

$$z_j^* = \underset{t=1,\ldots,T}{\operatorname{argmax}}\, d_{(t,j)}(x) \quad (2)$$

be which classifier has the highest confidence for class j. We select w^* via

$$w^* = \underset{j=1,\ldots,C}{\operatorname{argmax}}\, z_j^*. \quad (3)$$

Sum fusion is based on adding the support from each classifier. The confidence from each classifier for each prediction is added, and the prediction with the highest total is chosen. Let

$$s_j^* = \sum_{t=1,\ldots,T} d_{(t,j)}(x) \quad (4)$$

be the aggregate confidence in class j across all classifiers. Accordingly, w^* is determined via

$$w^* = \underset{j=1,\ldots,C}{\operatorname{argmax}}\, s_j^*. \quad (5)$$

b) Fuzzy aggregation: Fuzzy set theory includes a wide range of computational methods for fusing different types of information. One such operator is the fuzzy integral [16]. An advantage of the fuzzy integral, over other operators, is that it is a generator function and it produces many of the well-known crisp and fuzzy operators, e.g., ordered weighted averaging (OWA). The specific operator that the fuzzy integral is, depends on selection of the fuzzy measure (g). Herein, we explore the fuzzy integral to subsume and improve on the aforementioned operators. The discrete ChI is

$$C_g^j(\mathbf{d}) = \sum_{t=1}^{T} d_{((t),j)}(x)[g(A_t) - g(A_{t-1})], \quad (6)$$

where $C_g^j(\mathbf{d})$ is the integral for class j and fuzzy measure g, the inputs are sorted in decreasing order ($d_{((1),j)}(x) \geq \ldots \geq d_{((T),j)}(x)$), and A_t is the set of inputs from (1) to (t). Herein, like sum fusion, we use the ChI to fuse our classifier inputs per class, then the highest confident class is selected. Specifically, the ChI is learned using the quadratic programming-based convex optimization algorithm proposed in [17].

However, the fuzzy integral, for T inputs, has 2^T number of parameters (the fuzzy measure). This exponential number of parameters can grow quickly. In contexts like data-driven learning, this can be a limiting factor. The reader can refer to [17] and [18] for recent work on combating low variety datasets relative to the fuzzy integral and machine learning. Herein, we also explore the OWA, a subset of the ChI, as a tradeoff (performance versus simplicity). The OWA only has T parameters (the OWA weights). Specifically, we use a data driven and extended version of the OWA, called the linear order statistic neuron (LOSN) [19]. The reader can refer to [19] for full details about gradient descent-based LOSN optimization. Like the ChI, we use the LOSN to fuse across classifiers and the winning class is the one with the largest LOSN aggregate value. Diagrams for both the LOSN and ChI are shown in Fig. 3.

c) Unified networks: We also experiment with training a unified classifier where the feature vectors of the different classifiers are fused within the neural network training. Through the training process, the trainable parameters of the fusion network should be modified to accurately predict the classes. We test the fusion networks to compare their results to the other fusion methods and see if combination rules or fuzzy aggregation can improve upon the unified network. The spectral-spatial unified network (SSUN) combines the spectral LSTM and the spatial MSCNN to make predictions based only on HSI. For the SSUN, the combined framework is trained together, including concatenating the feature vectors of the two classifiers. We also test a second fusion network, the multimodal unified network (MUN) which combines the MSCNN and the LSTM that are trained on HSI along with the CNN that is trained on LiDAR. Each of the last FC layers from the CNN, LSTM, and MSCNN networks are concatenated into one FC layer and a single classifier is trained to accomplish multimodal classification.

IV. Experiments and Results

Experiments were performed with the LiDAR and HSI from 2018 GRSS dataset. As we did not have access to the test set from the contest, we used a random subset of the training data for testing. We subsequently split a portion of grss_dfc_2018 into training and validation data. The division and number of train and test samples is shown in Table 1. The entire subset of data used spatially separated the training and testing samples. We adopted 5-fold cross-validation, where the original dataset is partitioned into five equal size subsamples. For each run, one of the five subsamples is retained as the validation set to test the trained models, and four subsamples are used as the training data. We repeat this process of training the models five times, where each of the subsample sets are used as validation exactly once. We then average their results to produce a single estimation, shown in Table 2. The HSI had a ground sampling distance (GSD) of 1m while the LiDAR and corresponding labeled ground truth had a GSD of 0.5m, so we up-sampled the HSI to match the GSD. The HSI was processed for the spectral LSTM by band grouping to reduce the number of bands from 48 to five. For the spatial processing of the HSI and LiDAR, the data was divided into nine by nine pixel squares where the class of the center pixel determined the class label. The different methods were quantitatively compared through the overall accuracy, Cohen's kappa, and average accuracy. The overall accuracy depicts which portion of samples are classified correctly, and the average accuracy tells us the average accuracy per class. This average accuracy is the sum of each accuracy for each class predicted divided by the number of classes. This average accuracy value gives more emphasis to classes with fewer samples. Cohen's kappa is used as a robust statistic to test interrater or intrarater reliability. Note that a kappa value less than perfect (1.0) is a measure of both the agreement and disagreement among the raters.

TABLE I
SUBSET SAMPLES FROM GRSS_DFC_2018

Class	**Train**	**Test**
Unclassified	0	0
Healthy grass	9505	2394
Stressed grass	31189	7782
Artificial turf	656	153
Evergreen trees	13110	3223
Deciduous trees	4775	1220
Bare earth	4380	1017
Water	264	69
Residential buildings	37921	9487
Non-residential buildings	213994	53691
Roads	44397	11041
Sidewalks	32708	8080
Crosswalks	1518	356
Major thoroughfares	44722	11260
Highways	9283	2330
Railways	6673	1612
Paved parking lots	10983	2803
Unpaved parking lots	129	29
Cars	6415	1606
Trains	5174	1307
Stadium seats	6517	1619
Total:	484313	121079

A. Single classifier experiments

Each single-sensor based classifier was tested to give baseline results without any fusion. The MSCNN was tested on all training data, and had an overall classification accuracy of 96.610%. The LSTM achieved an overall classification accuracy of 95.922%. The LiDAR CNN had the worst performance among the single classifiers, with an overall classification accuracy of 80.536%. These are the base classifiers that are used for the combination rules, and fuzzy aggregation frameworks.

B. Unified Networks

Testing the unified networks required training a single framework that connects the final layers of multiple classifiers on different data types together. The SSUN, which is made up of the LSTM and MSCNN to combine spatial and spectral features of HSI, achieved an overall classification accuracy of 96.773%. The MUN, which also incorporates LiDAR, performed slightly better with an overall classification accuracy of

TABLE II
CLASSIFIERS AND RESULTS

Model	Spectral HSI	Spatial HSI	LiDAR	Overall Accuracy	Cohen's Kappa	Average Accuracy
Single modality classifiers						
LSTM	✓			95.922%	0.9470	94.569%
MSCNN		✓		96.610%	0.9559	94.497%
CNN			✓	80.536%	0.7438	68.061%
Unified networks						
SSUN	✓	✓		96.773%	0.9581	94.590%
MUN	✓	✓	✓	96.810%	0.9585	93.939%
Combination rules						
Majority voting	✓	✓	✓	97.420%	0.9664	95.681%
Maximum	✓	✓	✓	97.711%	0.9702	96.066%
Sum	✓	✓	✓	98.245%	**0.9772**	96.646%
Fuzzy aggregation						
LOSN	✓	✓	✓	**98.246%**	**0.9772**	96.681%
CHI	✓	✓	✓	98.135%	0.9758	**96.889%**

96.810%. The results of both the SSUN and MUN show that sensor fusion of multiple data types incrementally improved the results over each single data type.

C. Combination Rules

To test the performance of the combination rules, the MSCNN was trained on the spatial data, the LSTM was trained on the spectral data, and the CNN was trained on the LiDAR. The predicted classes and the confidence of those predictions were combined with algebraic combination rules. The combination rules, which were simple voting, sum, and maximum, performed better than the single classifiers or unified networks, with overall classification accuracies of 97.420%, 97.711%, and 98.245%, respectively. The most effective combination rule is sum, in which each classifier's confidence is summed to and highest value is taken.

D. Fuzzy Aggregation

To perform fuzzy aggregation, the training data is sent back through the three single-sensor based classifiers, and the confidences predicted by the model are used to train the aggregators to optimize the classification accuracy. The multiple classifiers used in these tests include the MSCNN, LSTM, and CNN. Both fuzzy aggregation methods perform well. The ChI has an overall classification accuracy of 98.135%, and the LOSN has the best performance with an overall classification accuracy of 98.246%. This is slightly better than the accuracy of the sum decision rule. It should be noted that the ChI achieves the best average accuracy of 96.889%, which is based on the accuracy per class. This is valuable because the dataset is not balanced, and there are not the same number of train and test samples for each class.

V. CONCLUSIONS

We develop a framework for multimodal sensor fusion of HSI and LiDAR and test different fusion methods within the framework, demonstrating that sensor fusion improves classification accuracy over any single modality. The tested fusion methods include established aggregation methods such as combination rules, unified neural networks, and fuzzy aggregation. The linear order statistic neuron (LOSN) has the best overall accuracy and the Choquet integral (ChI) has the best average accuracy, but the benefits of fuzzy aggregation extend beyond classification accuracy.

The success of the both the LOSN and ChI at the decision level of our deep learning framework suggests the promise of incorporating fuzzy neurons throughout neural networks. The use of fuzzy logic also shows promise for classification of challenging samples. Each pixel of HSI can contain the mixed spectra of multiple materials. Fuzzy measures can find a sample's membership in each different class instead of assigning it to a single class. Further, fuzzy aggregation can deal with incomplete information from each modality.

A major challenge in the implementation machine learning is the need for trusted and explainable machine learning models. Fuzzy aggregation provides a learned but transparent method for aggregation. Analyzing the values of the learned weights makes it possible to determine which type of aggregation it learns. In this case, the LOSN learns the sum combination rule, confirming that the sum combination rule is the best for this application. As the weights of the LOSN are interpretable, the LOSN can be considered a type of explainable artificial intelligence (AI).

While fuzzy aggregation methods lead to the best performance in overall accuracy and average accuracy, it also leads to other benefits when used within machine learning frameworks. Regardless of improvements in classification accuracy, between explainability and advantages with difficult samples, fuzzy aggregation has advantages over other fusion methods.

ACKNOWLEDGMENT

We would like to thank Derek T. Anderson and Charlie Veal at the Mizzou INformation and Data FUsion Laboratory (MINDFUL) at the University of Missouri for their help in the development, understanding and implementing the fuzzy aggregation tested in this paper. We would also like to thank Laura Hiatt and Sam Blisard at the U.S. Naval Research Laboratory (NRL) for their help in writing this paper.

REFERENCES

[1] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: an overview of methods, challenges, and prospects," Proceedings of the IEEE, vol. 103, no. 9, pp. 1449–1477, 2015.

[2] R. Polikar, "Ensemble based systems in decision making," in IEEE Circuits and Systems Magazine, vol. 6, no. 3, pp. 21-45, Third Quarter 2006.

[3] 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," in IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 9, pp. 1638-1642, Sept. 2017.

[4] J. A. Benediktsson and P. Ghamisi, Spectral-spatial classification of hyperspectral remote sensing images. Boston: Artech House, 2015.

[5] P. Goel, S. Prasher, R. Patel, J. Landry, R. Bonnell, and A. Viau, "Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn," Computers and Electronics in Agriculture, vol. 39, no. 2, pp. 67–93, 2003.

[6] M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 114, pp. 24–31, 2016.

[7] G. Camps-Valls, and L. Bruzzone, "Kernel-based methods for hyperspectral image classification." IEEE Transactions on Geoscience and Remote Sensing 43.6 (2005): 1351-1362.

[8] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi and J. A. Benediktsson, "Deep Learning for Hyperspectral Image Classification: An Overview," in IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 9, pp. 6690-6709, Sept. 2019.

[9] C. Debes et al., "Hyperspectral and LiDAR Data Fusion: Outcome of the 2013 GRSS Data Fusion Contest," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 6, pp. 2405-2418, June 2014.

[10] Y. Xu, B. Du, L. Zhang, D. Cerra, M. Pato, E. Carmona, S. Prasad, N. Yokoya, R. Hansch, and B. L. Saux, "Advanced Multi-Sensor Optical Remote Sensing for Urban Land Use and Land Cover Classification: Outcome of the 2018 IEEE GRSS Data Fusion Contest," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 6, pp. 1709–1724, 2019

[11] 2018 IEEE GRSS Data Fusion Contest. Online: http://www.grss-ieee.org/community/technical-committees/data-fusion".

[12] B. Bigdeli, F. Samadzadegan, and P. Reinartz, "Fusion of hyperspectral and LIDAR data using decision template-based fuzzy multiple classifier system," International Journal of Applied Earth Observation and Geoinformation, vol. 38, pp. 309–320, 2015.

[13] Y. Xu, L. Zhang, B. Du and F. Zhang, "Spectral–Spatial Unified Networks for Hyperspectral Image Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 10, pp. 5893-5909, Oct. 2018.

[14] L. Mou, P. Ghamisi and X. X. Zhu, "Deep Recurrent Neural Networks for Hyperspectral Image Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 7, pp. 3639-3655, July 2017.

[15] H. Luo, "Shorten Spatial-spectral RNN with Parallel-GRU for Hyperspectral Image Classification," arXiv.org, 2018. [Online]. Available: https://arxiv.org/abs/1810.12563.

[16] J. M. Keller, D. B. Fogel, and D. Liu, Fundamentals of computational intelligence: neural networks, fuzzy systems and evolutionary computation. Hoboken: John Wiley & Sons, 2016.

[17] M. A. Islam, D. T. Anderson, A. J. Pinar, and T. C. Havens, "Data-Driven Compression and Efficient Learning of the Choquet Integral," IEEE Transactions on Fuzzy Systems, vol. 26, no. 4, pp. 1908–1922, 2018.

[18] M. A. Islam, D. T. Anderson, A. Pinar, T. C. Havens, G. Scott, and J. M. Keller, "Enabling Explainable Fusion in Deep Learning with Fuzzy Integral Neural Networks," IEEE Transactions on Fuzzy Systems, pp. 1–1, 2019.

[19] C. Veal et al., "Linear Order Statistic Neuron," 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), New Orleans, LA, USA, 2019, pp. 1-6.