Open Set Recognition

Prof. Jefersson A. dos Santos

jefersson@dcc.ufmg.br

PATREO

DCC

DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

UFMG

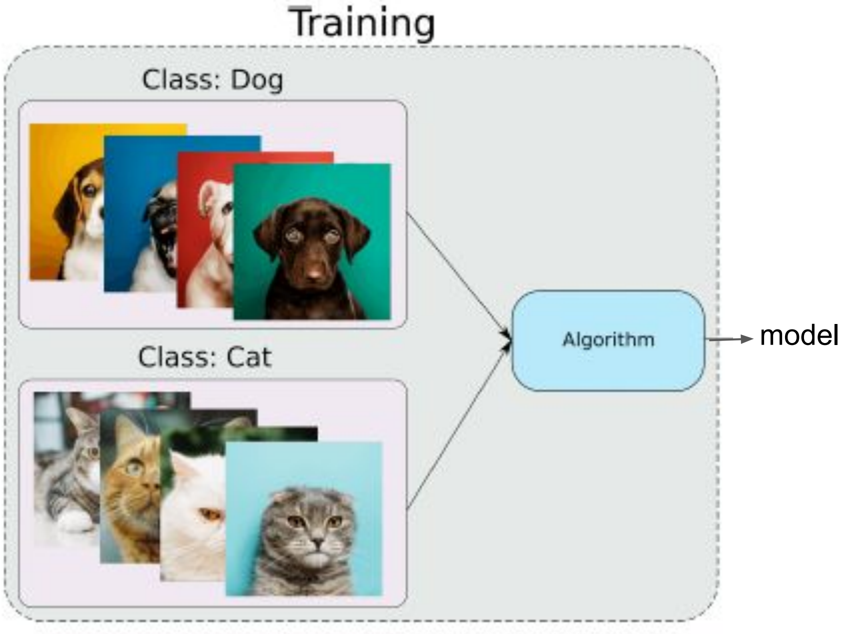
Outline

1. Introduction
2. Classification/OSR Approaches
    a. Shallow Methods
    b. Deep-based Methods
3. Segmentation Approaches
    a. Pixel-wise
    b. Fully Convolutional
4. Final Remarks

Introduction

Supervised Classification

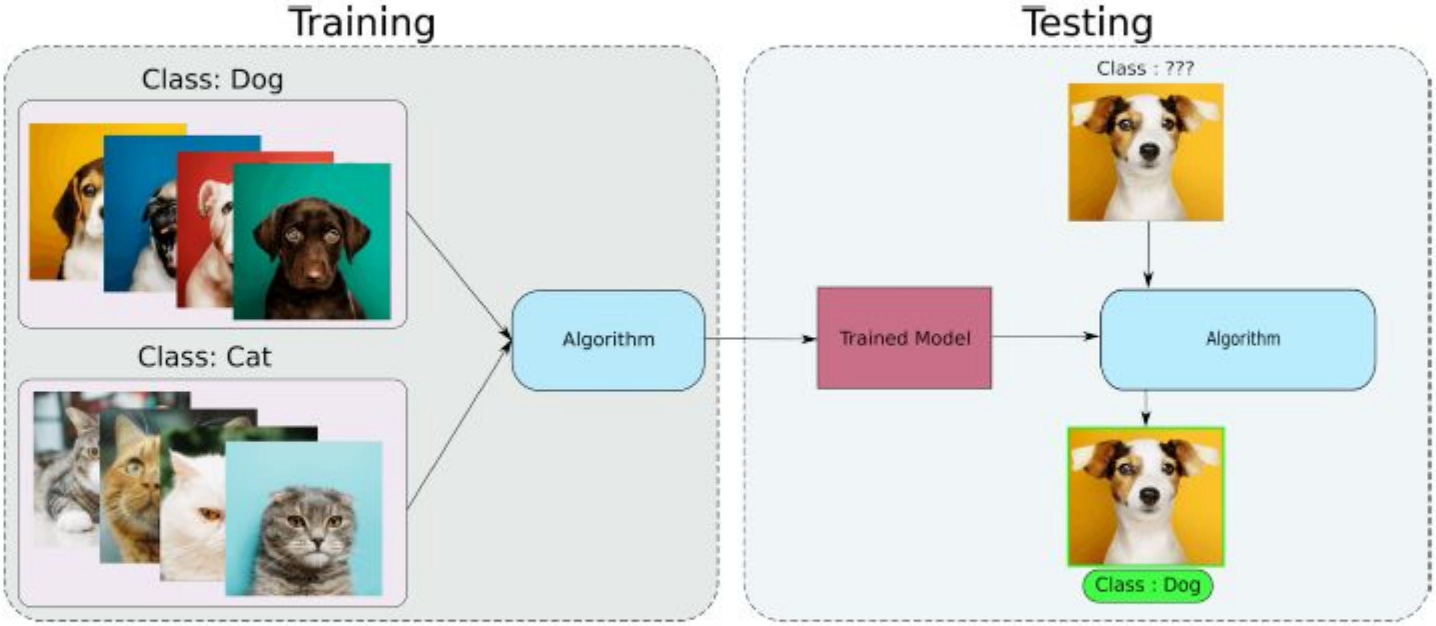
Closed Set



In closed-set scenarios, training and testing have the same classes and feature spaces, a factor that facilitates the work of supervised models.

Supervised Classification

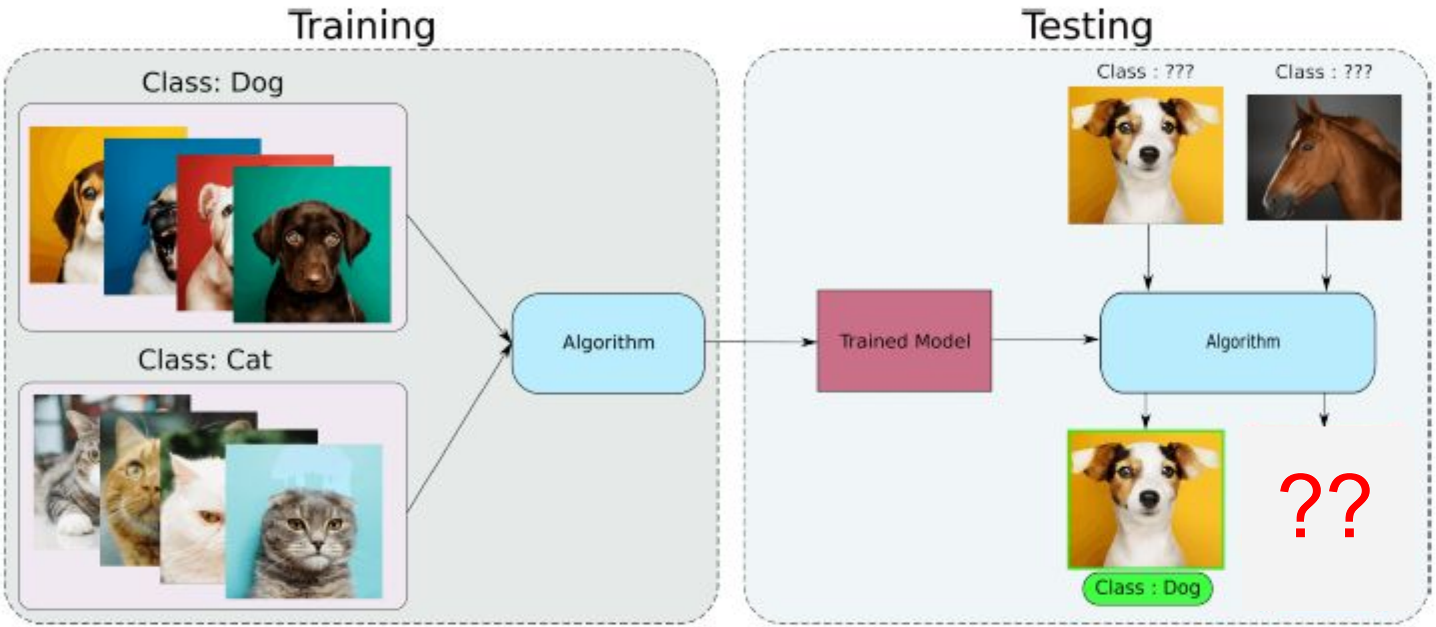
Closed Set



In closed-set scenarios, training and testing have the same classes and feature spaces, a factor that facilitates the work of supervised models.

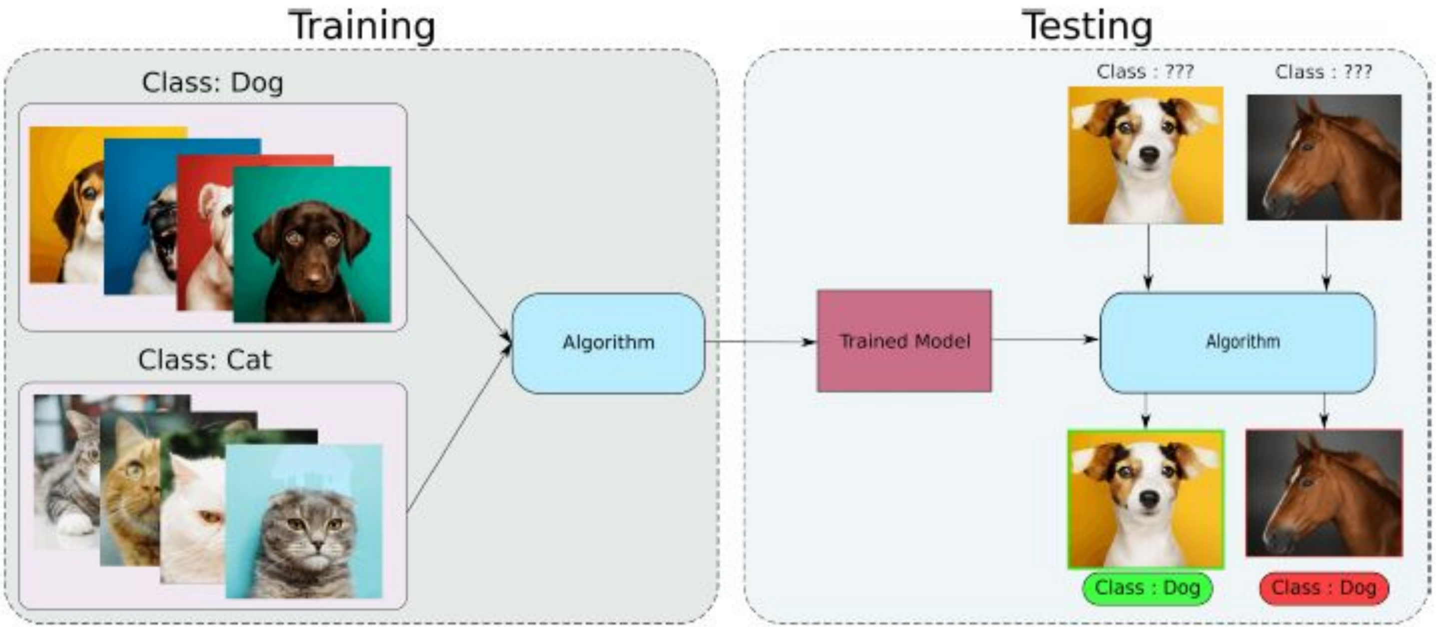
Supervised Classification

Open Set



In realistic scenarios, unknown situations may appear, a factor that impairs the robustness of the supervised models.

Supervised Classification
Open Set

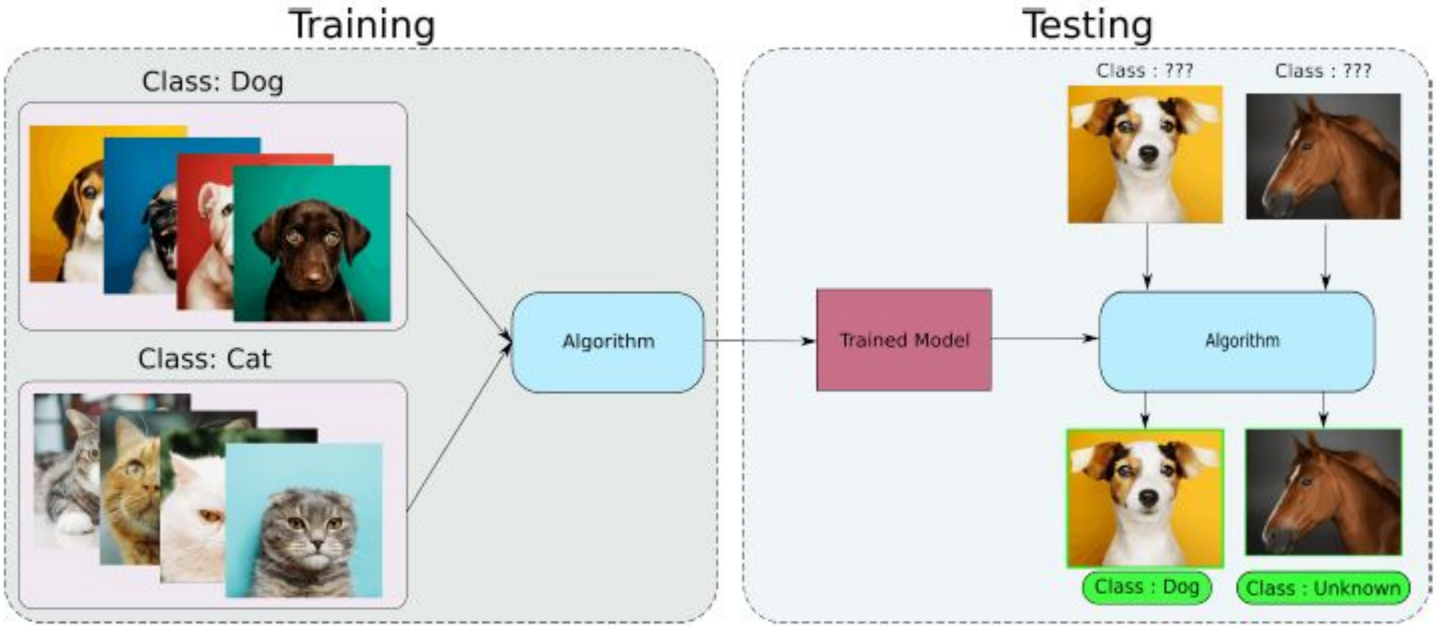


In realistic scenarios, unknown situations may appear, a factor that impairs the robustness of the supervised models.

WRONG!

Supervised Classification

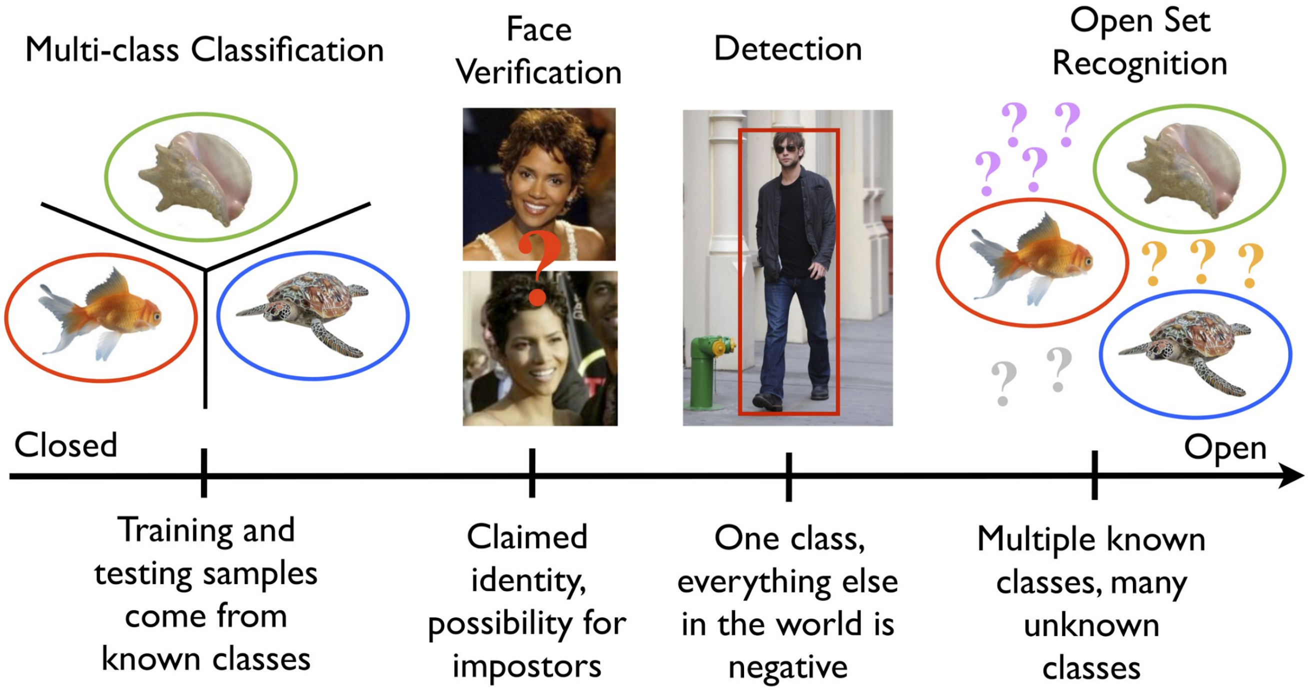
Open Set



\o/

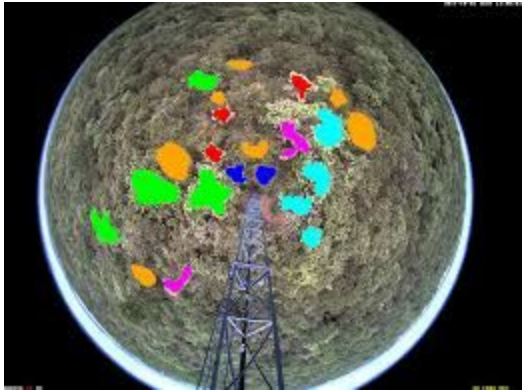
Open Set Recognition

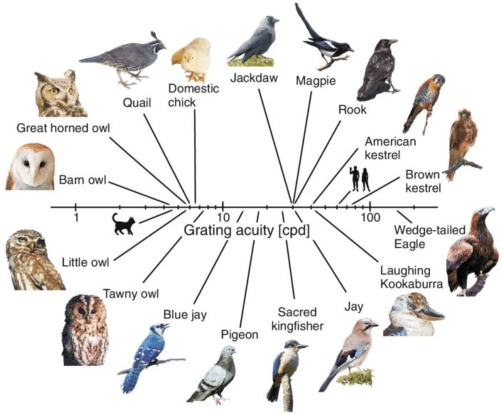
https://www.wjscheirer.com/projects/openset-recognition/

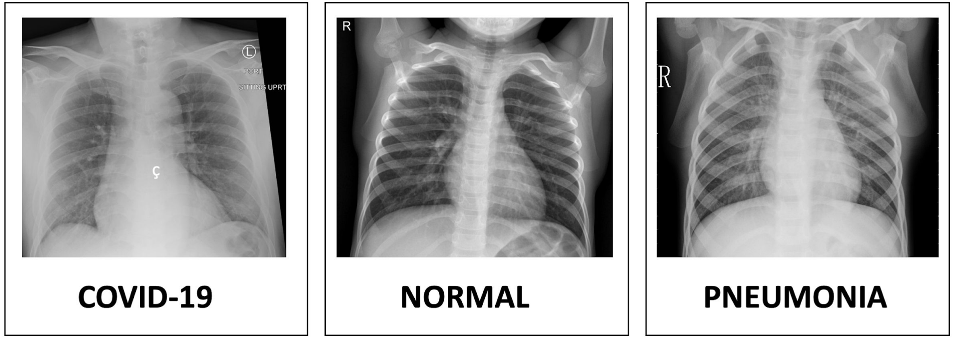


Scheirer, W. J., de Rezende Rocha, A., Sapkota, A., & Boult, T. E. (2012). Toward open set recognition. *IEEE transactions on pattern analysis and machine intelligence*, *35*(7), 1757-1772.

OSR in the Wild



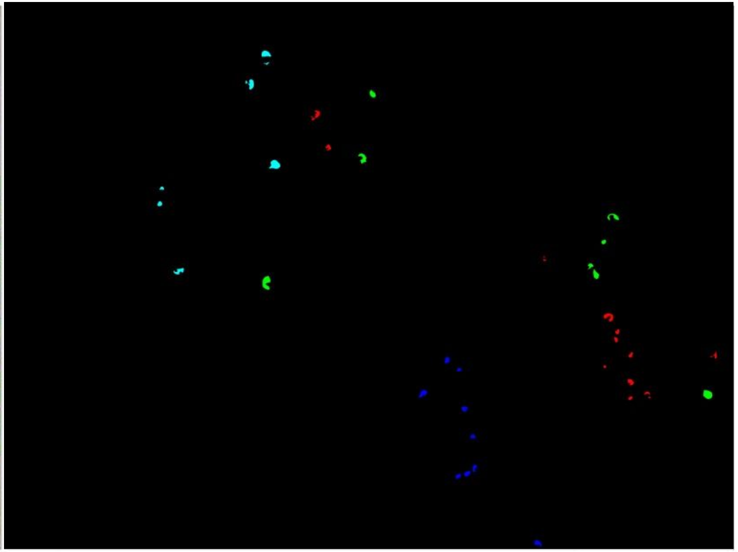




OSR in the Wild

Serra do Cipó, State of Minas Gerais

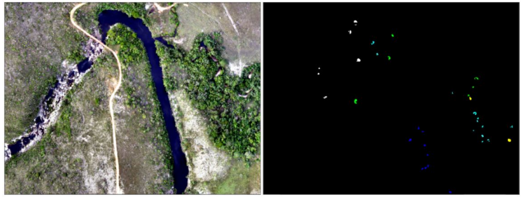
UAV Image



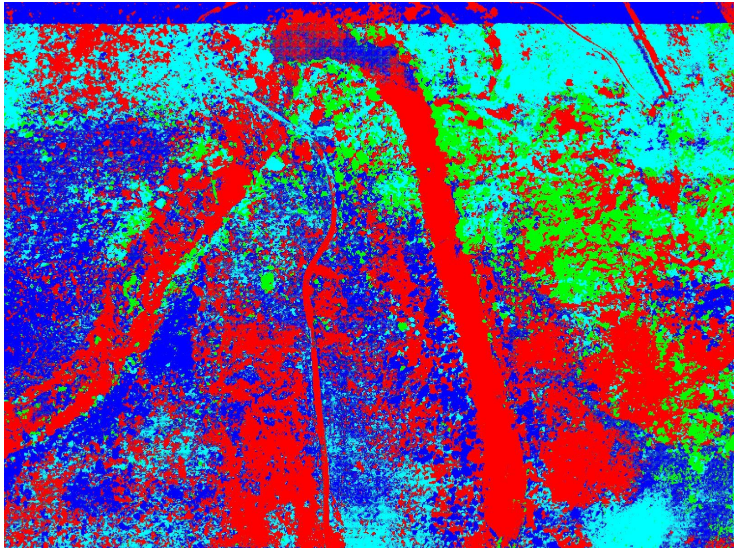
Endemic Species

Nogueira, K., dos Santos, J. A., Menini, N., Silva, T. S., Morellato, L. P. C., & Torres, R. D. S. (2019). Spatio-Temporal Vegetation Pixel Classification by Using Convolutional Networks. *IEEE Geoscience and Remote Sensing Letters*, *16*(10), 1665-1669.

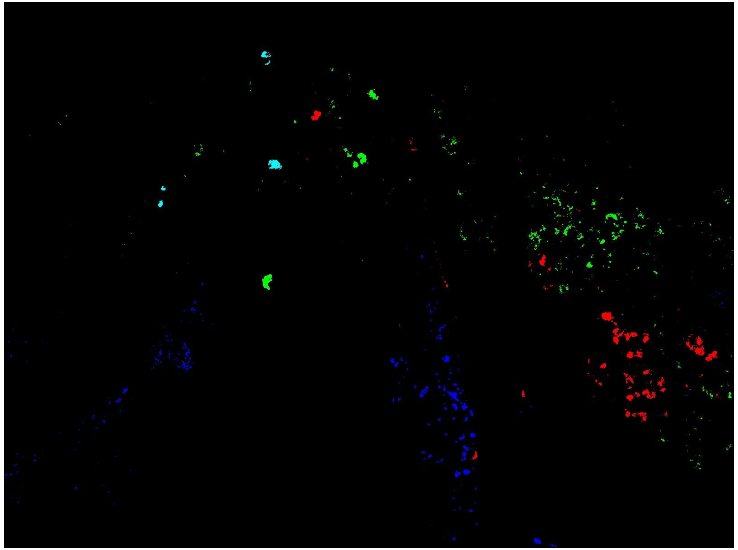
OSR in the Wild



Serra do Cipó, State of Minas Gerais



Closed Set



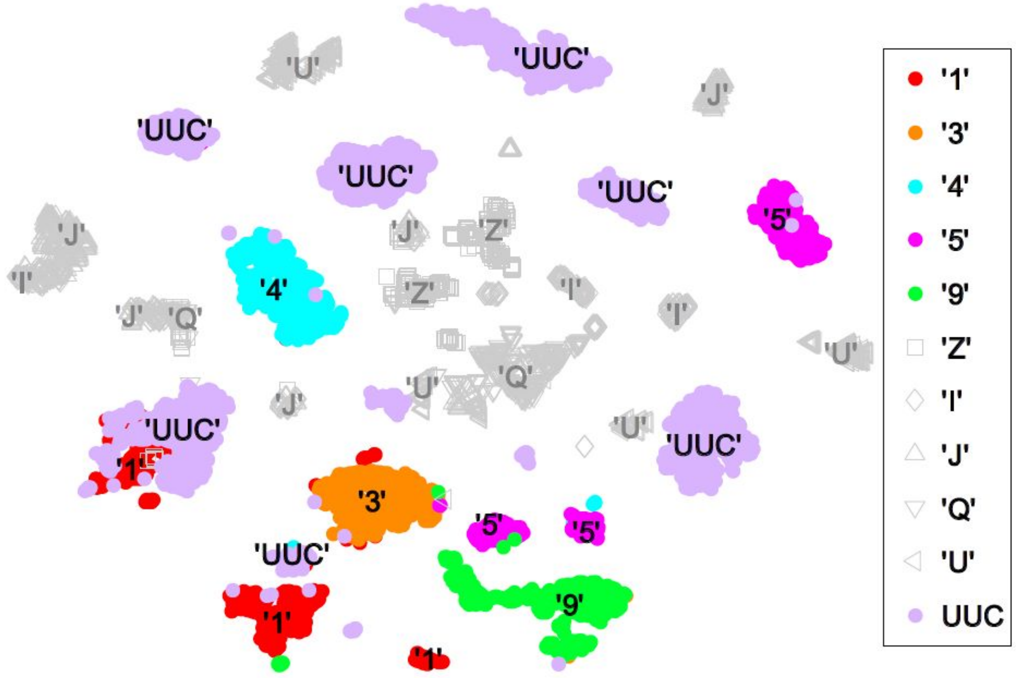
Open Set

Basic recognition categories of classes

- Known known classes (KKCs):
  - Classes with positive/negative samples and labels
- Known unknown classes (KUCs):
  - Classes with negative labels that were not grouped in significative classes (e.g. background classes)
- Unknown known classes (UKCs):
  - Classes with no available samples in training, but available side information (e.g. metadata or attribute information)
- Unknown unknown classes (UUCs):
  - Classes without any information during training

Geng, C., Huang, S.-J., & Chen, S. (2020). Recent Advances in Open Set Recognition: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Basic recognition categories of classes



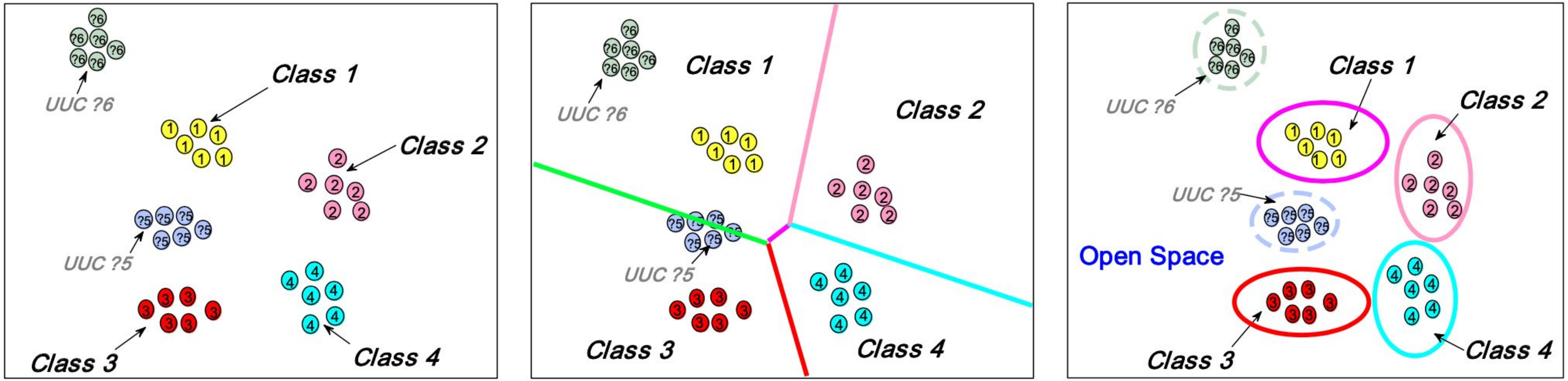
An example of visualizing KKCs, KUCs, and UUCs from the real data distribution using t-SNE.

- **KKCs:** ’1’,’3’,’4’,’5’,’9’ are randomly selected from PENDIGITS
- **UUCs:** remaining classes from PENDIGITS
- **KUCs:** ’Z’,’I’,’J’,’Q’,’U’ randomly selected from LETTER

This visualization also indicates that the distribution of one class may consist of multiple subclass/subclusters, e.g., class ’1’, ’5’, ’U’, etc.

Open Set Recognition (OSR)

- OSR describes a scenario where new classes (UUCs) not seen in training appear during testing.
    - the model must correctly classify known entries (KKCs) and identify UUC entries.



(a) Distribution of the original data set. (b) Traditional recognition/classification problem. (c) Open set recognition/classification problem.

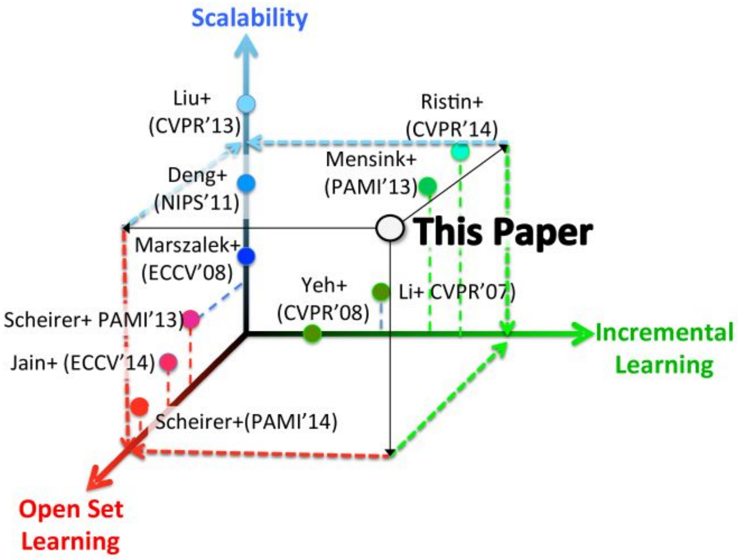
Open Set Recognition Vs Related Tasks

SETTING / TASK	TRAINING	TESTING	GOAL
Traditional Classification	Known known classes	Known known classes	Classifying known known classes
Classification with Reject Option	Known known classes	Known known classes	Classifying known known classes & rejecting samples of low confidence
One-class Classification (Anomaly Detection)	Known known classes & few or none outliers from KUCs	Known known classes & few or none outliers	Detecting outliers
One/Few-shot Learning	Known known classes & a limited number of UKCs' samples	Unknown known classes	Identifying unknown known classes
Generalized Few-shot Learning	Known known classes & a limited number of UKCs' samples	Known known classes & unknown known classes	Identifying known known classes & unknown known classes
Zero-shot Learning	Known known classes & side-information[1]	Unknown known classes	Identifying unknown known classes
Generalized Zero-shot Learning	Known known classes & side-information[1]	Known known classes & unknown known classes	Identifying known known classes & unknown known classes
Open Set Recognition	Known known classes	Known known classes & unknown unknown classes	Identifying known known classes & rejecting unknown unknown classes
Generalized Open Set Recognition	Known known classes & side-information[2]	Known known classes & Unknown unknown classes	Identifying known known classes & cognizing unknown unknown classes

Open World

"A recognition system in the 'open world' has to continuously update with additional object categories and be robust to unseen categories and have minimum downtime".

(Bendale, 2015)

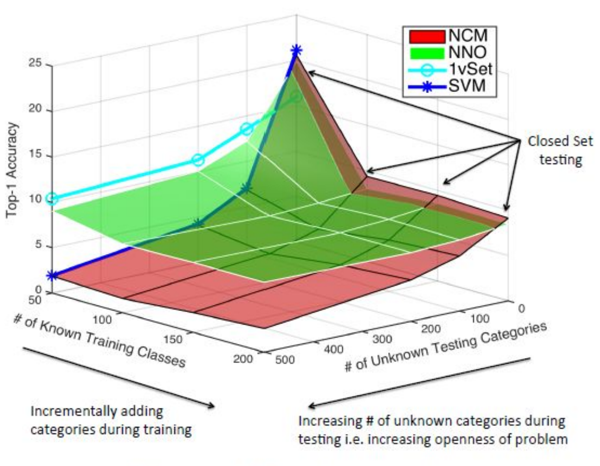


Bendale, A., & Boult, T. (2015). Towards open world recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1893-1902).

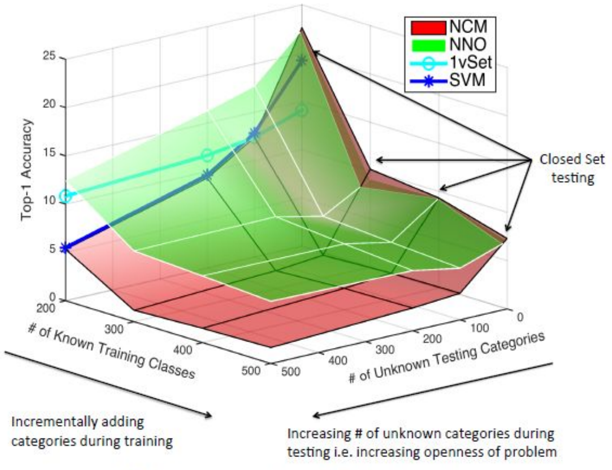
Open World

"A recognition system in the 'open world' has to continuously update with additional object categories and be robust to unseen categories and have minimum downtime".

(Bendale, 2015)



(a) 50 categories in metric learning phase.



(b) 200 categories in metric learning phase.

Bendale, A., & Boult, T. (2015). Towards open world recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1893-1902).

Open-Set vs Few-Shot Learning

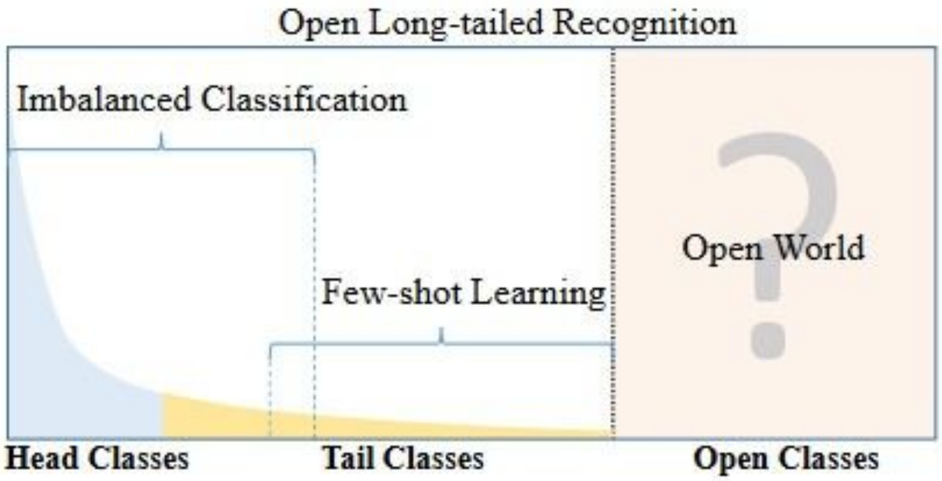


Figure 1: Our task of open long-tailed recognition must learn from long-tail distributed training data in an open world and deal with imbalanced classification, few-shot learning, and open-set recognition over the entire spectrum.

Liu, Z., Miao, Z., Zhan, X., Wang, J., Gong, B., & Yu, S. X. (2019). Large-scale long-tailed recognition in an open world. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2537-2546).

Openness

- The number of unknown classes is a critical point in the process.
- **Openness** is defined by:
  - C_{TR}: Classes used in training.
  - C_{TA}: Classes to be recognized.
  - C_{TE}: Classes used in the test.

$$O = 1 - \sqrt{\frac{2 \times |C_{\mathrm{TR}}|}{|C_{\mathrm{TA}}| + |C_{\mathrm{TE}}|}}$$

- The greater the Openness the more 'open' the problem is.
- Openness equal to 0 is a completely closed problem.

$$C_{\mathrm{TA}} \subseteq C_{\mathrm{TR}} \subseteq C_{\mathrm{TE}}$$

Openness

- CTA must be a subgroup of CTR.
- For a specific problem, Openness depends only on KKCs CTR and UUCs CTE. With that, O*:

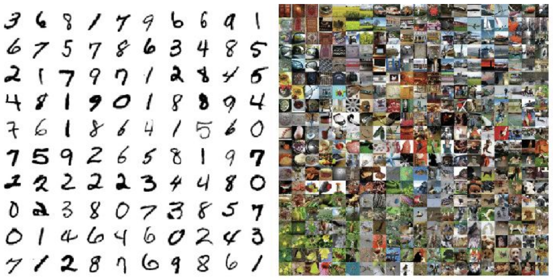
$$O^* = 1 - \sqrt{\frac{2 \times |C_{\mathrm{TR}}|}{|C_{\mathrm{TR}}| + |C_{\mathrm{TE}}|}}$$

Experimental Setup for Open-Set Recognition Tasks

Datasets

- Multiclass class datasets are usually used, but the classes are randomly divided into KKC (training and testing) and UUC (testing).

Dataset	Dados	Classes	KKC	UUC
LETTER	20000	26 (769 dados with 16 features)	10	16
PENDIGITS	10992	10 (1099 dados with 16 features)	5	5
COIL20	1440	20 (72 imagens 16x16)	10	10
YALEB	2414	38 (64 imagens 32x32)	10	28
MNIST	65000~	10 (6500~ imagens 128x128)	6	4
SVHN	100000~	10 (10000~ imagens 32x32)	6	4
CIFAR10	6000	10 (600 imagens 32x32)	6	4
TINY-IMAGENET	110000	200 (550 imagens 32x32)	20	180



Evaluation Measures

- Metrics need to be adapted to consider the recognition of UUCs.
- Notations:
  - TPi: True Positive; TNi: True Negative; FPi: False Positive; FNi: False Negative; for the KKC i.
  - TU: Correct UUCs; FU: wrongly rejected.

Accuracy:

$$\mathcal{A} = \frac{\sum_{i=1}^{C}(TP_i + TN_i)}{\sum_{i=1}^{C}(TP_i + TN_i + FP_i + FN_i)}.$$

Closed Set

$$\mathcal{A}_O = \frac{\sum_{i=1}^{C}(TP_i + TN_i) + TU}{\sum_{i=1}^{C}(TP_i + TN_i + FP_i + FN_i) + (TU + FU)}.$$

Open Set

Important: It does not define OSR methods very well because of the imbalance between KKC and UUC.

Evaluation Measures

Normalized Accuracy:

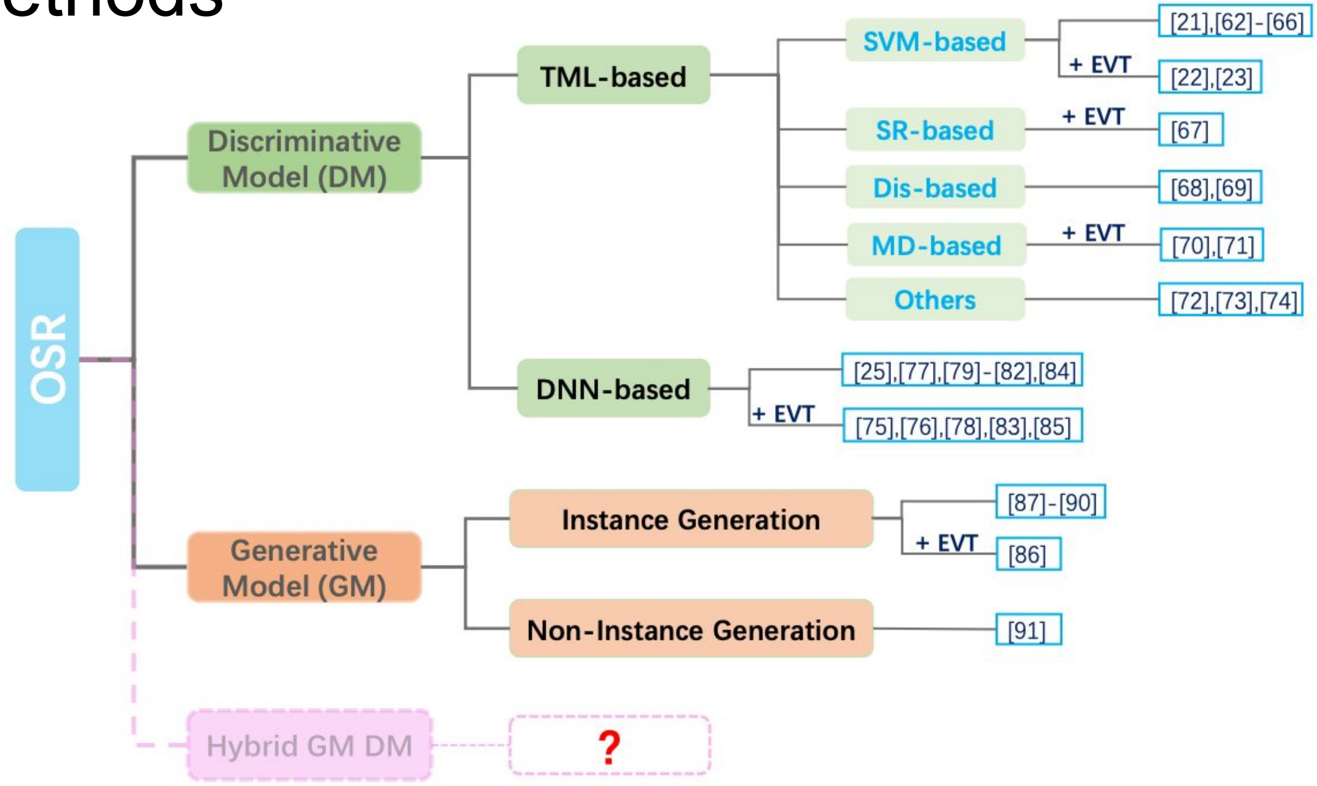
$$\mathrm{NA} = \lambda_r \mathrm{AKS} + (1 - \lambda_r)\mathrm{AUS}.$$

$$\mathrm{AKS} = \frac{\sum_{i=1}^{C}(TP_i + TN_i)}{\sum_{i=1}^{C}(TP_i + TN_i + FP_i + FN_i)}, \ \mathrm{AUS} = \frac{TU}{TU + FU}$$

- It considers different weights for KKC and UUC
  - $0<\lambda_r<1$

Open-Set Recognition Methods

OSR methods



Geng, C., Huang, S.-J., & Chen, S. (2020). Recent Advances in Open Set Recognition: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1–1.

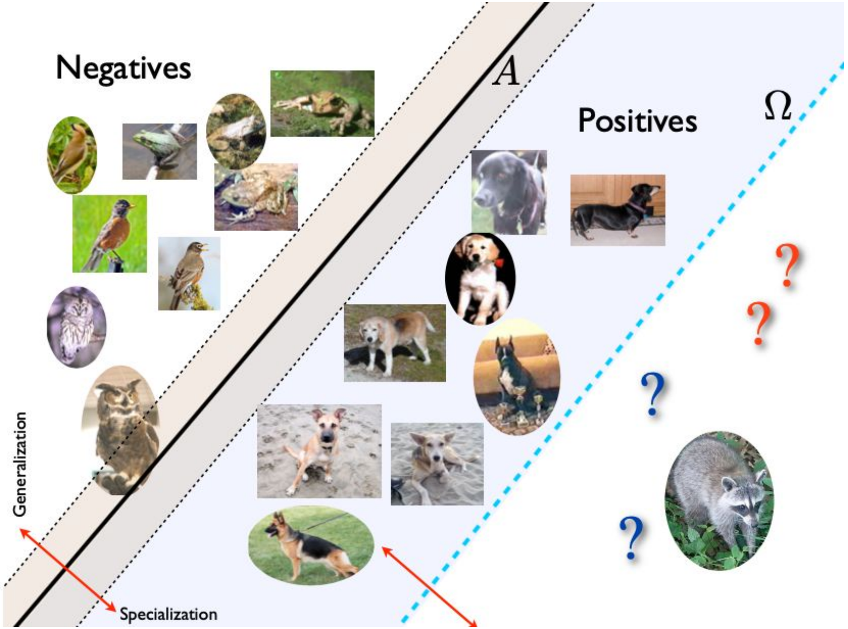
Open-Set Recognition Methods

TML-Based

TML-Based

1-vs-Set Machine

- It incorporates a risk term to take into account the space beyond the KKCs classes.
- It seeks to minimize risk (reduce generalization) considering margin restrictions (avoid over-specialization).
- Adds another hyperplane, parallel to the SVM plane defining the class boundaries.
- Objects between the 2 planes belong to the class.
- **Limitation:** It only separates one class.



Scheirer, W. J., de Rezende Rocha, A., Sapkota, A., & Boult, T. E. (2012). Toward open set recognition. *IEEE transactions on pattern analysis and machine intelligence*, *35*(7), 1757-1772.

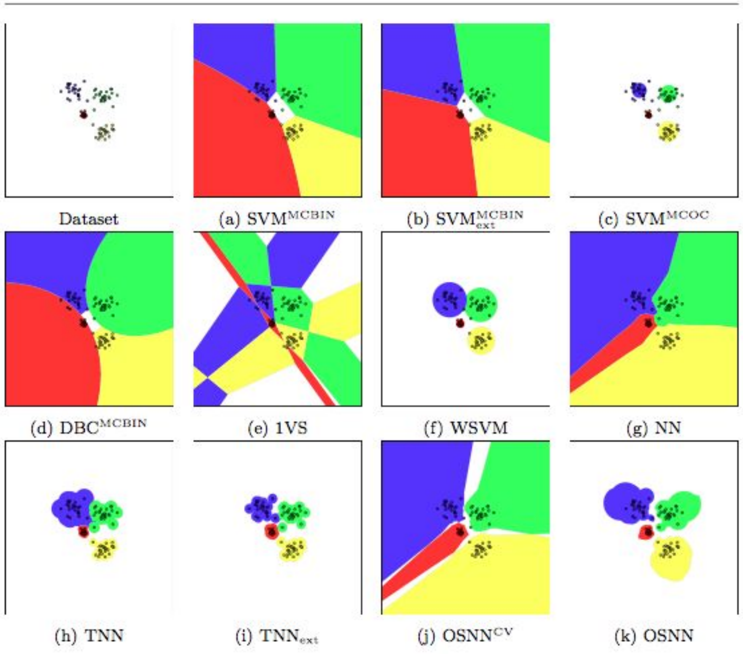
TML-Based

Open Set Nearest Neighbor (OSNN)

- It applies a decision limit in the ratio of similarity between the two most similar classes.
- It find the nearest neighbors u and t from an example of s tests and calculate the ratio between Euclidean distances:

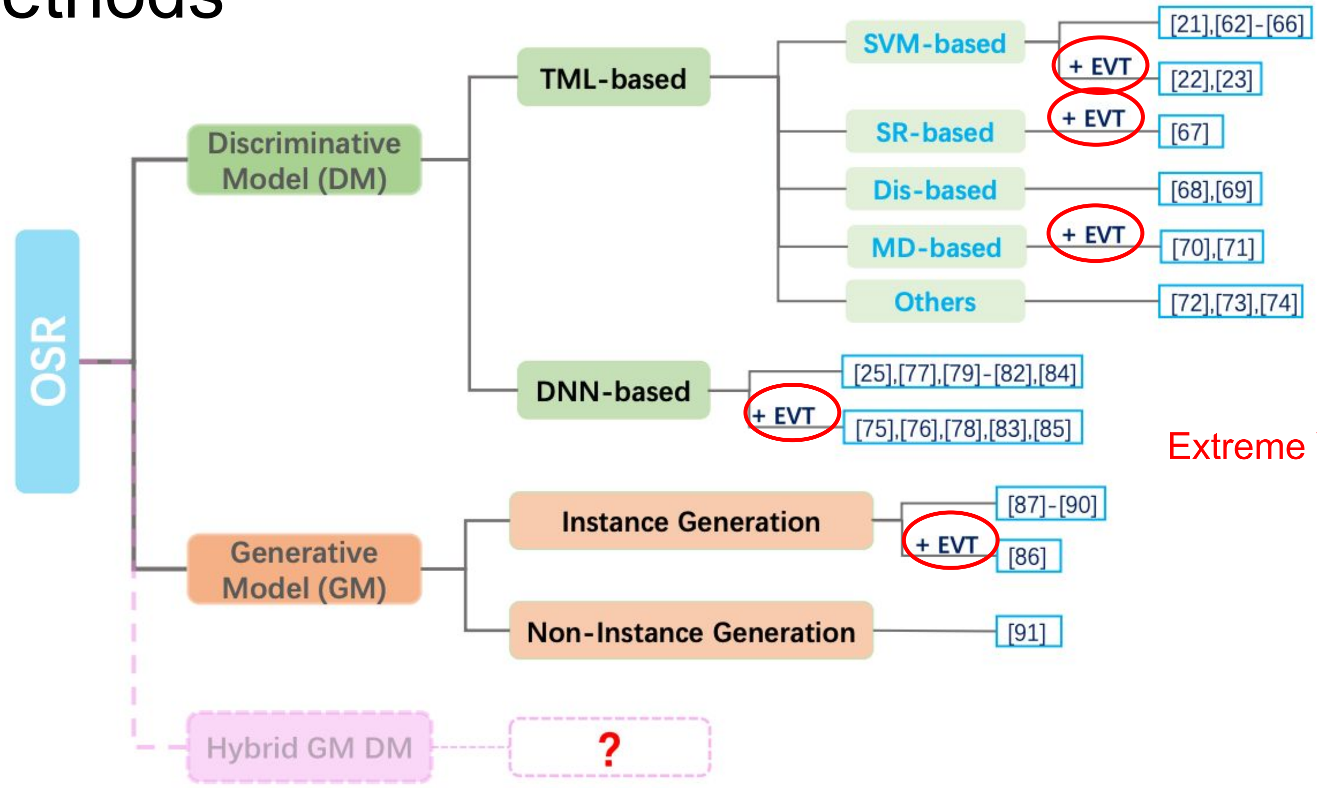
$$\text{Ratio} = d(s,t)/d(s,u)$$

- If the ratio is less than the limit, the entry is UUC.
    - Multiclass.



Mendes Jr, P. R., De Souza, R. M., Werneck, R. D. O., Stein, B. V., Pazinato, D. V., de Almeida, W. R., & Rocha, A. (2017). Nearest neighbors distance ratio open-set classifier. *Machine Learning*, *106*(3), 359-386.

OSR methods



Extreme Value Theory

Geng, C., Huang, S.-J., & Chen, S. (2020). Recent Advances in Open Set Recognition: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1–1.

Extreme Value Theory (EVT)

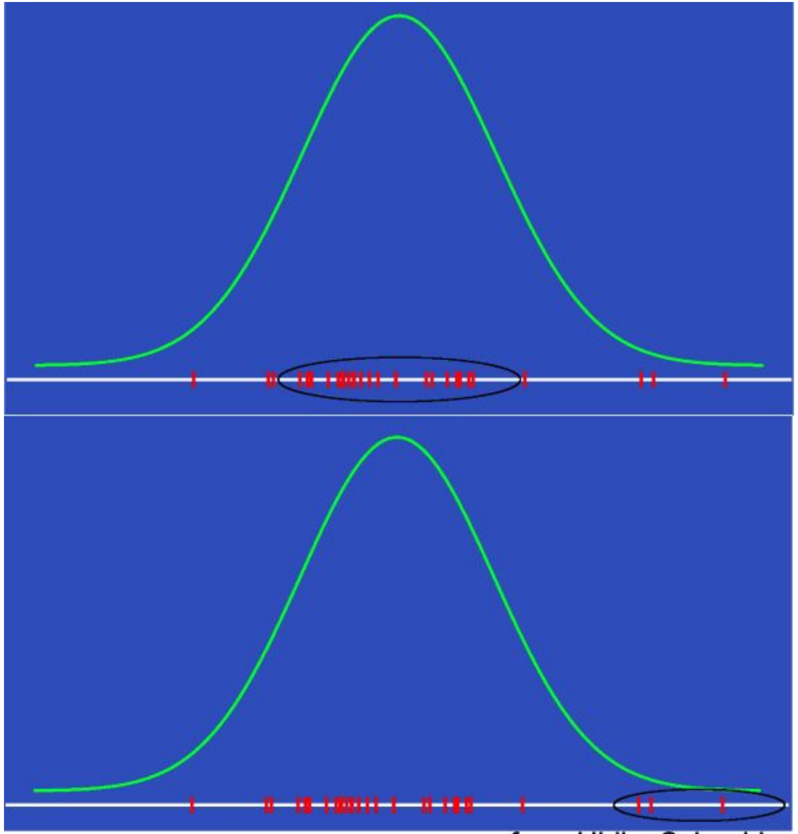
In classical statistics:

- focus on AVERAGE behavior of stochastic process
- Central Limit Theorem

Extreme Value Theory:

- focus on extreme and rare events
- Fisher-Tippett Theorem

What is an extreme?



from Ulrike Schneider

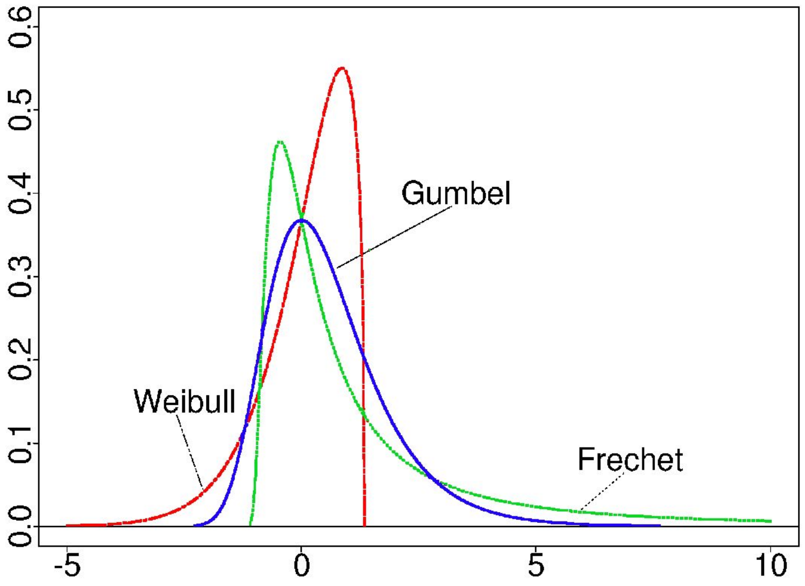
Coles, S., Bawa, J., Trenner, L., & Dorazio, P. (2001). *An introduction to statistical modeling of extreme values* (Vol. 208, p. 208). London: Springer.

Extreme Value Theory (EVT)

Extremes from a very large domain of stochastic processes follow **generalized extreme value (GEV) distribution:**

$$H(x) = \begin{cases} \exp \ -\left\{1 + \ v\left(\dfrac{x-u}{s}\right)\right\}^{-\frac{1}{v}}, v \ \neq 0 \\ \exp(-e^{x}) \qquad\qquad\qquad\qquad\quad , v = 0 \end{cases}$$

u – location
s – scale
v – shape

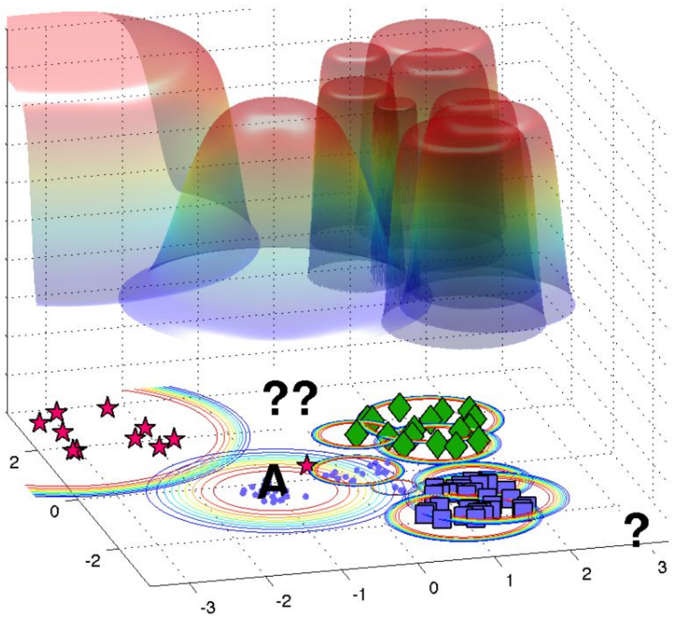


Shape parameter v defines several distributions:

- Gumbel: v = 0
- **Weibull: v < 0**
- Fréchet: v > 0

Extreme Value Machine (EVM)

- Each class is represented by a group of extreme vectors (EV) associated with a probability function to include examples.
- This function is modeled in relation to half distances relative to a reference point.
- It is based on margin distribution models by example.
- Following the margin distribution theorem, a point X is able to estimate the distance distribution itself.
- The probability is given by the density function Ψ.
- The prediction is made by the probability of the density function of the closest class with a threshold to separate UUCs.
- **Open World Recognition. Multiclass.**

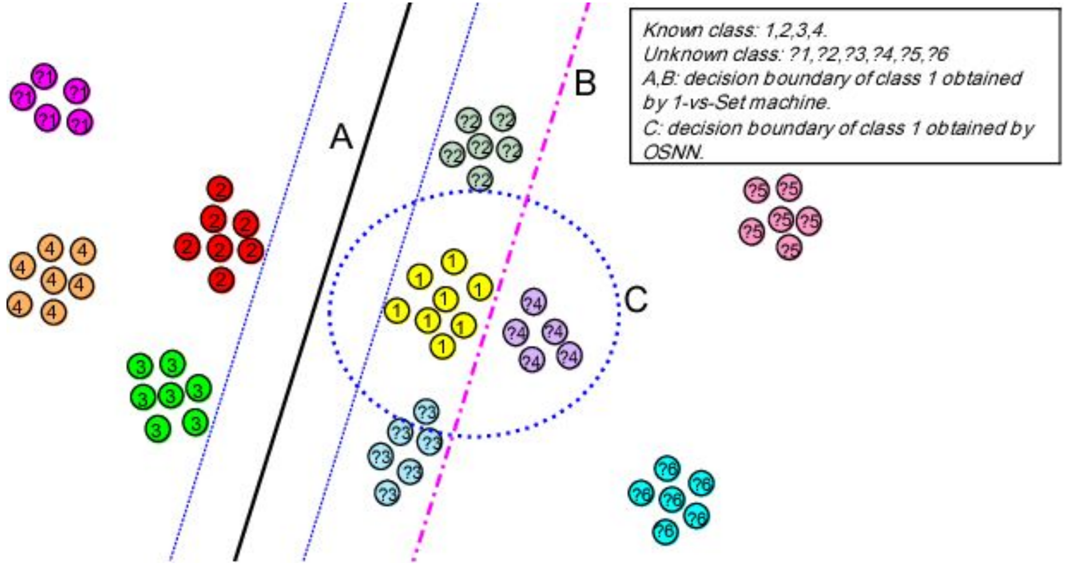


$$\Psi(x_i, x'; \kappa_i, \lambda_i,) = \exp^{-\left(\frac{||x_i - x'||}{\lambda_i}\right)^{\kappa_i}}$$

Rudd, E. M., Jain, L. P., Scheirer, W. J., & Boult, T. E. (2017). The extreme value machine. *IEEE transactions on pattern analysis and machine intelligence*, *40*(3), 762-768.

TML-Based

Collective Decision for OSR
CD-OSR



Proposed Solution:

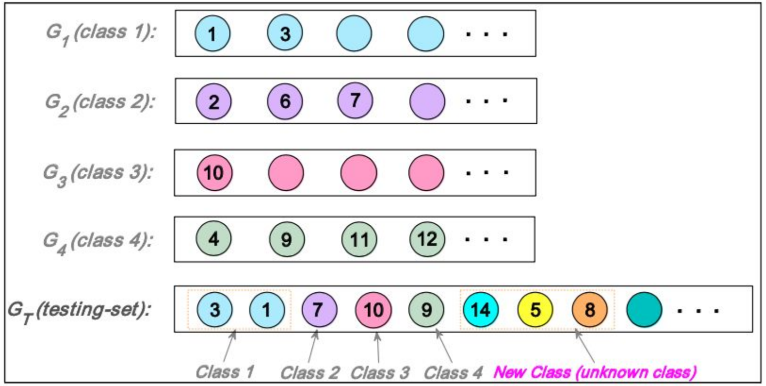
- Adaptation of the Hierarchical Dirichlet Process (HDP)
- It can address both the open set recognition and new class discovery simultaneously

Geng, C., & Chen, S. (2020). Collective decision for open set recognition. *IEEE Transactions on Knowledge and Data Engineering*.

TML-Based

Collective Decision for OSR
CD-OSR

- Training:
  - a co-clustering is done, where each class is modeled as a group in HDP using Gaussian mixtures (GMM) with an unknown number of components.
  - the set of parameters that each class represents is defined in this process
- Test:
  - a data / group is co-clustered using CD-OSR and each class is represented as a set of subclasses.
  - It is not necessary to define thresholds to define the limit between KKCs and UUCs.



It has several limitations, such as the need for co-clustering for each batch.

Geng, C., & Chen, S. (2020). Collective decision for open set recognition. *IEEE Transactions on Knowledge and Data Engineering*.

Experimental Setup
Shallow Learning

- Datasets: LETTER, PENDIGITS, COIL20, YALEB
- Metric: F-Measure
- Most methods adopt threshold-based strategies.
- The threshold is defined considering only KKC data.

Experimental Results

Shallow Learning

Dataset / Method		1-vs-Set	W-OSVM	W-SVM	P_I-SVM	SROSR	OSNN	EVM	CD-OSR
LETTER	O^*=0%	81.51±3.94	95.64±0.37	95.64±0.25	96.92±0.36	84.21±2.49	83.12±17.41	96.59±0.50	**96.94±1.36**
	O^*=15.48%	55.43±3.18	83.83±2.85	91.24±1.48	90.89±1.80	74.36±5.10	73.20±15.21	89.81±0.40	**91.51±1.58**
	O^*=25.46%	42.08±2.63	73.37±1.67	85.72±0.85	84.16±1.01	66.50±8.22	64.97±13.75	82.81±2.42	**86.21±1.46**
PENDIGITS	O^*=0%	97.17±0.58	94.84±1.46	98.82±0.26	**99.21±0.29**	97.43±0.93	98.55±0.71	98.42±0.73	99.16±0.25
	O^*=8.71%	78.43±1.93	87.22±1.71	93.05±1.85	92.38±2.68	96.33±1.59	95.55±1.30	96.97±1.37	**98.75±0.65**
	O^*=18.35%	61.29±2.52	78.55±4.91	88.39±3.14	87.60±4.78	93.53±3.26	90.11±4.15	92.88±2.79	**98.43±0.73**
COIL20	O^*=0%	89.59±1.81	93.94±1.87	86.83±1.82	89.30±1.45	97.12±0.60	79.61±7.41	97.68±0.88	**97.71±0.94**
	O^*=10.56%	70.21±1.67	90.82±2.31	85.64±2.47	87.68±2.02	96.68±0.32	73.01±6.18	95.69±1.46	**97.32±1.50**
	O^*=18.35%	57.72±1.50	87.97±5.40	84.54±3.79	86.22±3.34	**96.45±0.66**	66.18±4.49	93.62±3.33	95.12±2.14
YALEB	O^*=0%	87.99±2.42	82.60±3.54	86.01±2.42	**93.47±2.74**	88.09±3.41	81.81±8.40	68.94±6.47	89.75±1.15
	O^*=23.30%	49.36±1.96	63.43±5.33	84.56±2.19	**88.96±1.16**	83.99±4.19	72.90±9.41	54.40±5.77	88.00±2.19
	O^*=35.45%	34.37±1.44	55.40±5.26	83.44±2.02	**86.63±0.60**	81.38±5.26	67.24±7.29	46.64±5.40	85.56±1.07

Experimental Results - Conclusions
TML-based methods

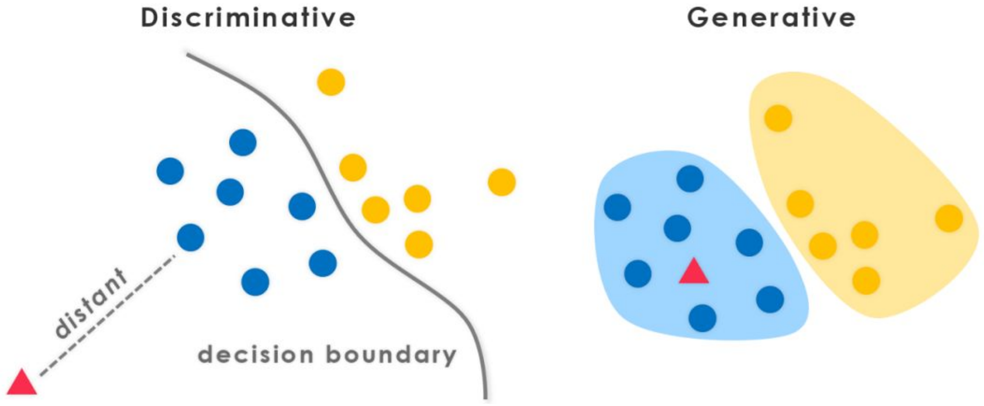
1. With the increase in Openness, threshold-based methods may experience reduced performance.
2. 1-vs-Set methods worsen performance with increasing Openness.
3. OSNN is dependent on the distribution of the characteristics of the dataset.
4. CD-OSR can better adapt to UUCs due to HDP and presents better performance in most datasets. However, HDP is difficult to apply to data with many dimensions and has a high computational cost.

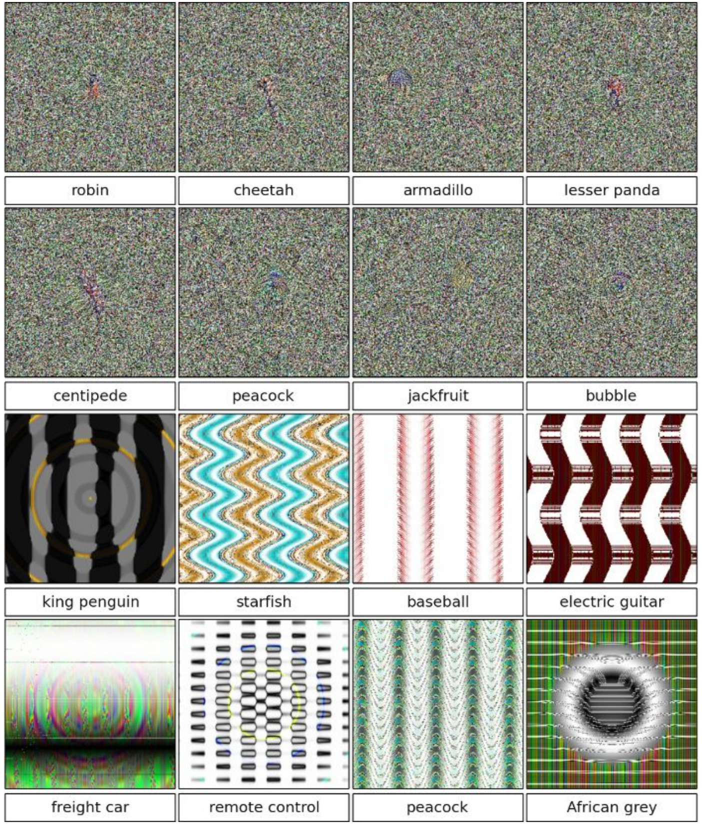
Open-Set Recognition Methods

DNN-Based

Deep Learning vs OSR

High confidence predictions for unrecognizable images

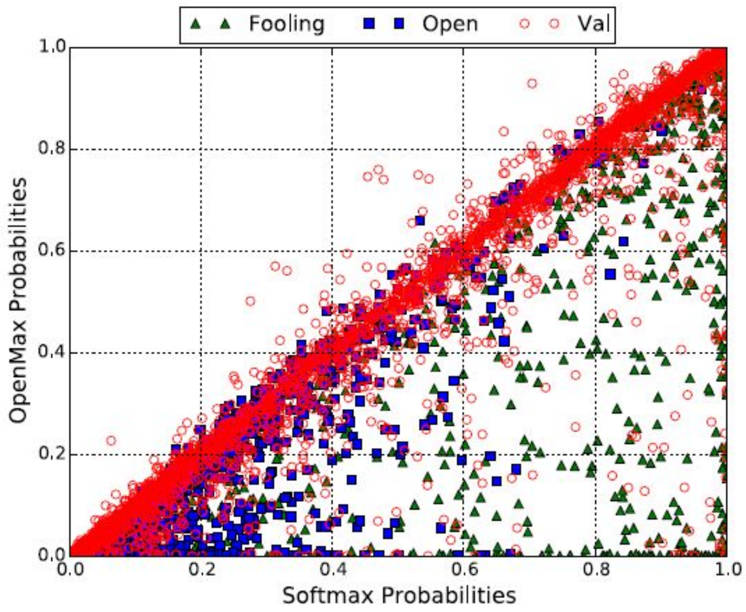




Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 427-436).

OpenMax

- Method extends softMax so that it is possible to compute UUCs.
- The network is trained with SoftMax.
- Each class is represented by a mean activation vector (MAV) of the penultimate layer of the network.
- Each class's MAVs are redistributed according to the **Weibull distribution** which is used to define the probability that a sample of data belongs to the class.



Bendale, A., & Boult, T. E. (2016). Towards open set deep networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1563-1572).

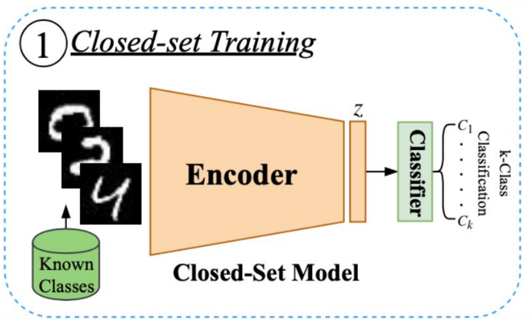
DNN-Based

Class Conditioned Auto-Encoder (C2AE)

Given images in a batch $\{X_1, X_2, ..., X_N\} \in \mathcal{K}$, and their corresponding labels $\{y_1, y_2, ..., y_N\}$. Here N is the batch size and $\forall y_i \in \{1, 2, .., k\}$. The encoder ($\mathcal{F}$) and the classifier ($\mathcal{C}$) with parameters Θ_f and Θ_c, respectively are trained using the following cross entropy loss,

$$\mathcal{L}_c(\{\Theta_f, \Theta_c\}) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{k} \mathbb{I}_{y_i}(j) \log[p_{y_i}(j)], \quad (1)$$

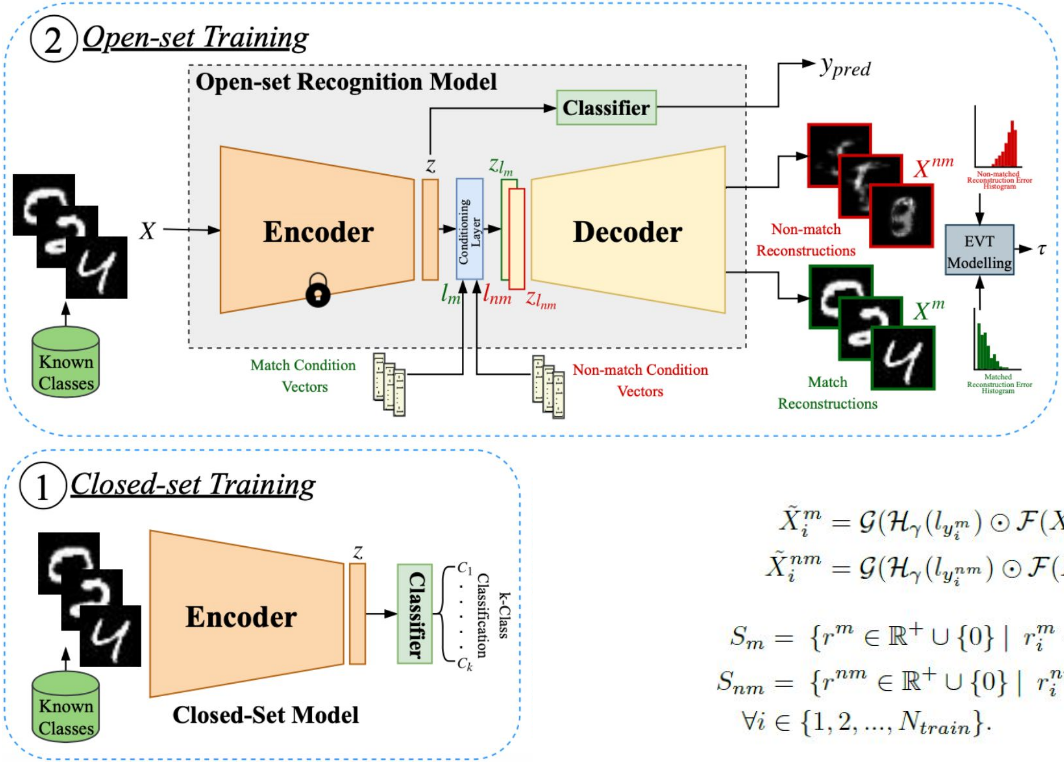
where, $\mathbb{I}_{y_i}$ is an indicator function for label y_i (*i.e.*, one hot encoded vector) and $p_{y_i} = \mathcal{C}(\mathcal{F}(X_i))$ is a predicted probability score vector. $p_{y_i}(j)$ is probability of the i^{th} sample being from the j^{th} class.



Oza, P., & Patel, V. M. (2019). C2ae: Class conditioned auto-encoder for open-set recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2307-2316).

DNN-Based

Class Conditioned Auto-Encoder (C2AE)



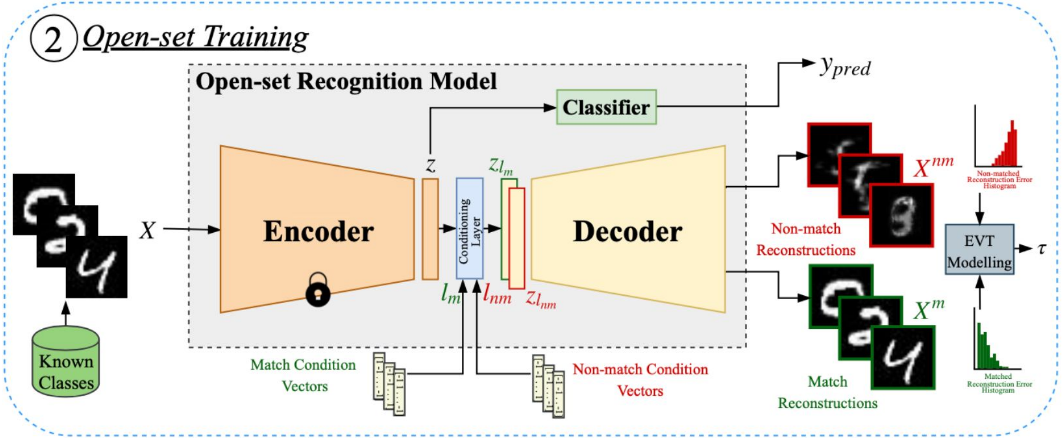
$$\tilde{X}_i^m = \mathcal{G}(\mathcal{H}_\gamma(l_{y_i^m}) \odot \mathcal{F}(X_i) + \mathcal{H}_\beta(l_{y_i^m})),$$
$$\tilde{X}_i^{nm} = \mathcal{G}(\mathcal{H}_\gamma(l_{y_i^{nm}}) \odot \mathcal{F}(X_i) + \mathcal{H}_\beta(l_{y_i^{nm}})),$$

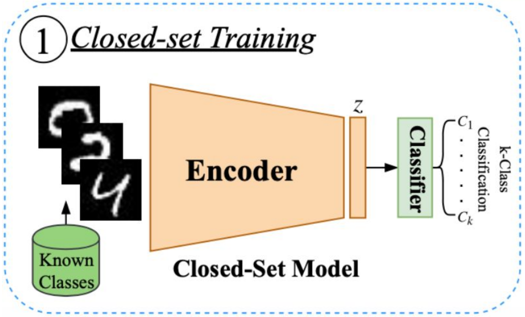
$$S_m = \{r^m \in \mathbb{R}^+ \cup \{0\} \mid r_i^m = \|X_i - \tilde{X}_i^m\|_1\},$$
$$S_{nm} = \{r^{nm} \in \mathbb{R}^+ \cup \{0\} \mid r_i^{nm} = \|X_i - \tilde{X}_i^{nm}\|_1\},$$
$$\forall i \in \{1, 2, ..., N_{train}\}.$$

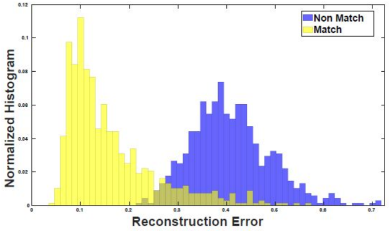
Oza, P., & Patel, V. M. (2019). C2ae: Class conditioned auto-encoder for open-set recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2307-2316).

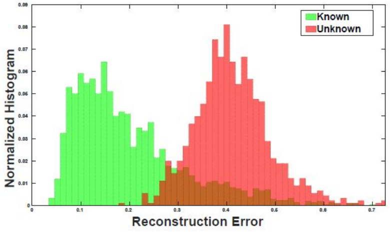
DNN-Based

Class Conditioned Auto-Encoder (C2AE)





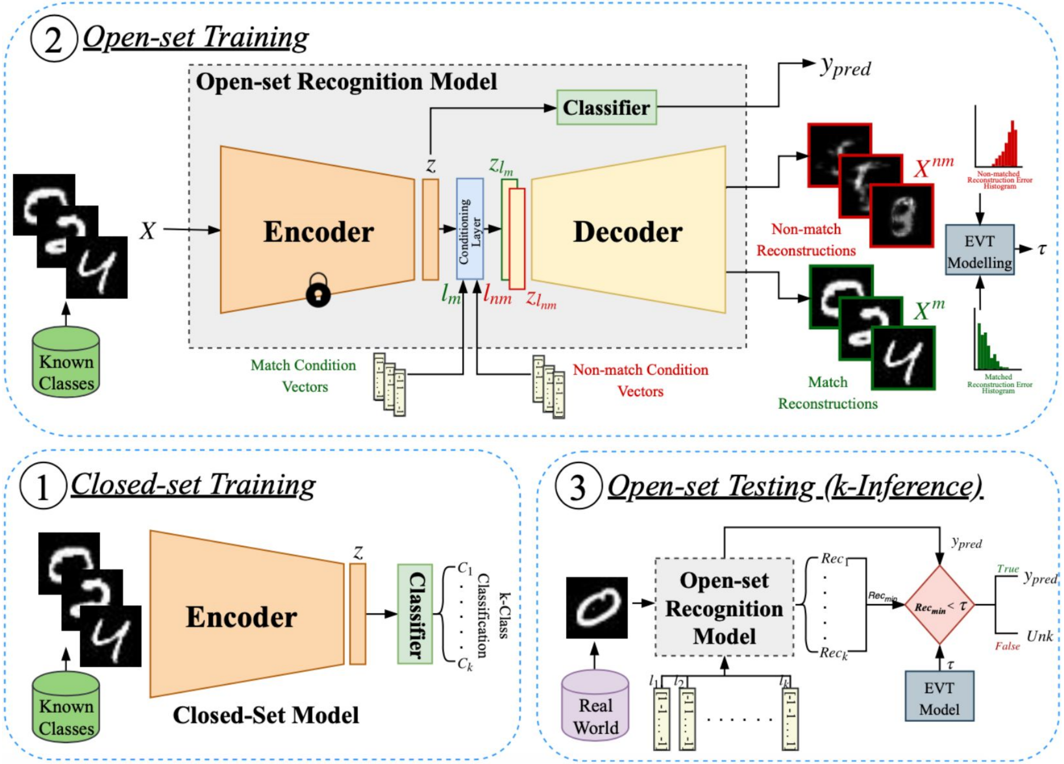




Oza, P., & Patel, V. M. (2019). C2ae: Class conditioned auto-encoder for open-set recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2307-2316).

DNN-Based

Class Conditioned Auto-Encoder (C2AE)



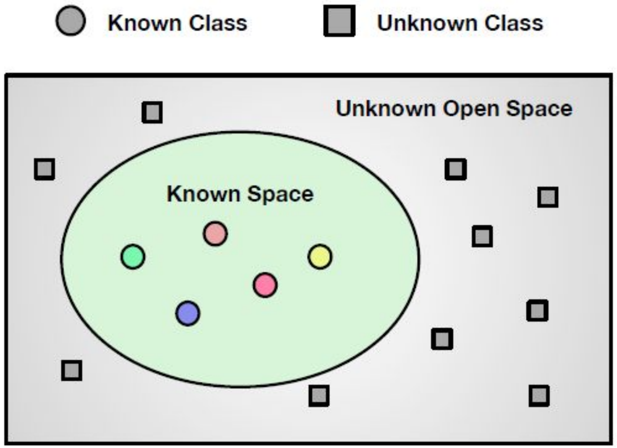
Oza, P., & Patel, V. M. (2019). C2ae: Class conditioned auto-encoder for open-set recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2307-2316).

Open-Set Recognition Methods

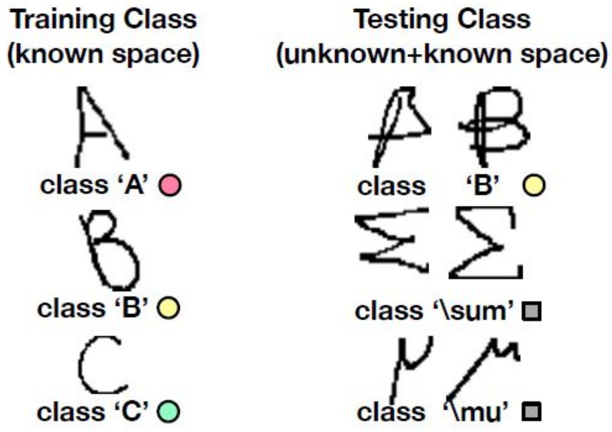
Generative Models

G-OpenMax

- Uses GAN to create UUCs.
- It makes possible to have an explicit probability of unknown data.



(a) Known Space vs. Open Space

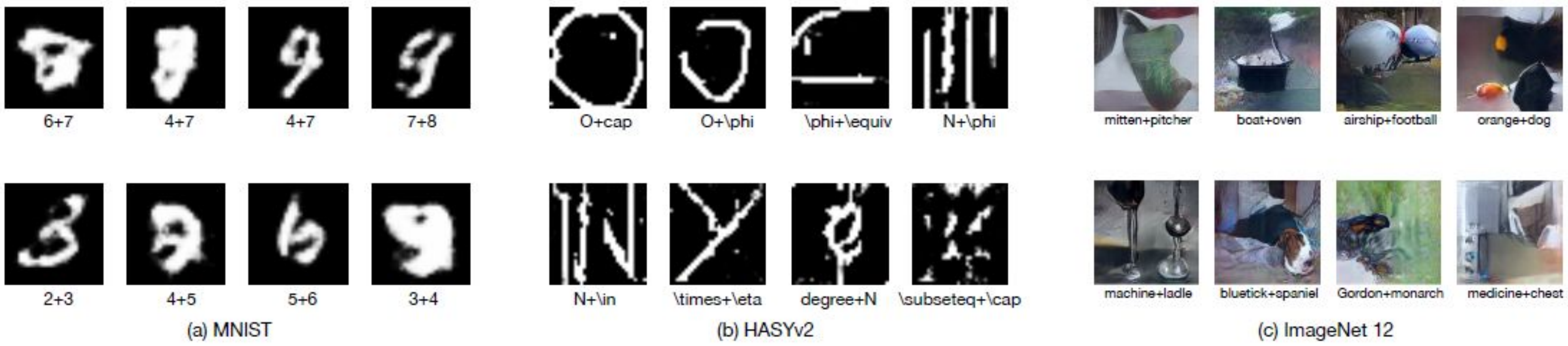


(b) Multi-class Open Set Classification

Ge, Z., Demyanov, S., Chen, Z., & Garnavi, R. (2017). Generative OpenMax for multi-class open set classification. In *British Machine Vision Conference 2017*. British Machine Vision Association and Society for Pattern Recognition.

G-OpenMax

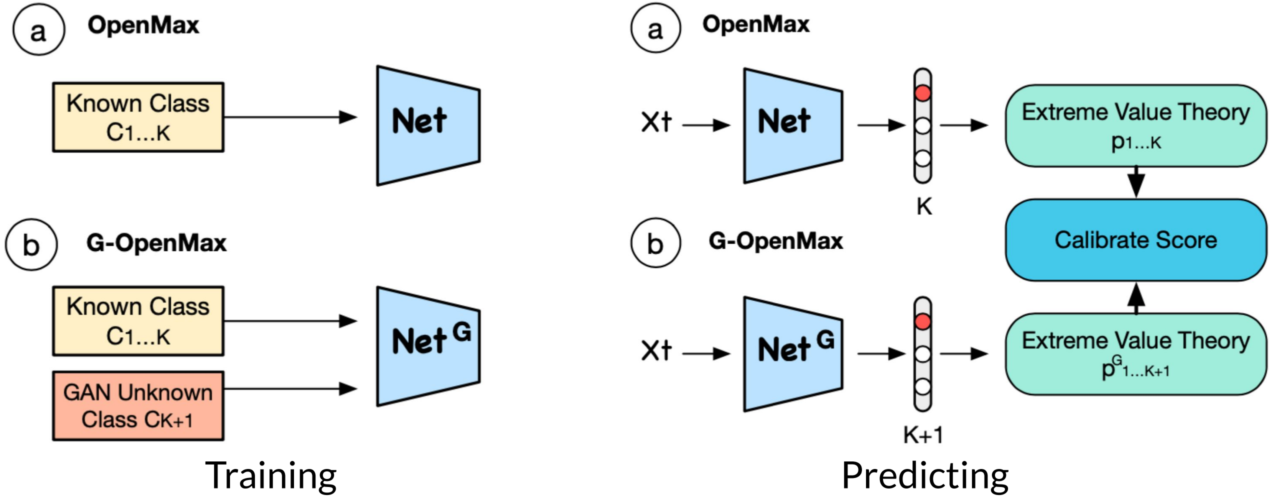
- Uses GAN to create UUCs.
- It makes possible to have an explicit probability of unknown data.



Ge, Z., Demyanov, S., Chen, Z., & Garnavi, R. (2017). Generative OpenMax for multi-class open set classification. In *British Machine Vision Conference 2017*. British Machine Vision Association and Society for Pattern Recognition.

G-OpenMax

- Uses GAN to create UUCs.
- It makes possible to have an explicit probability of unknown data.

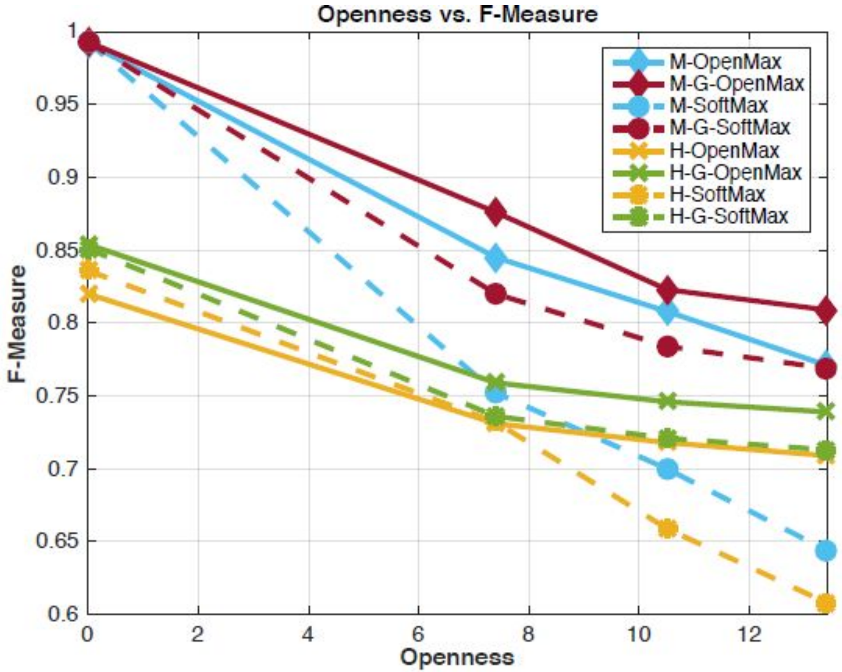


Ge, Z., Demyanov, S., Chen, Z., & Garnavi, R. (2017). Generative OpenMax for multi-class open set classification. In *British Machine Vision Conference 2017*. British Machine Vision Association and Society for Pattern Recognition.

G-OpenMax

Conclusions

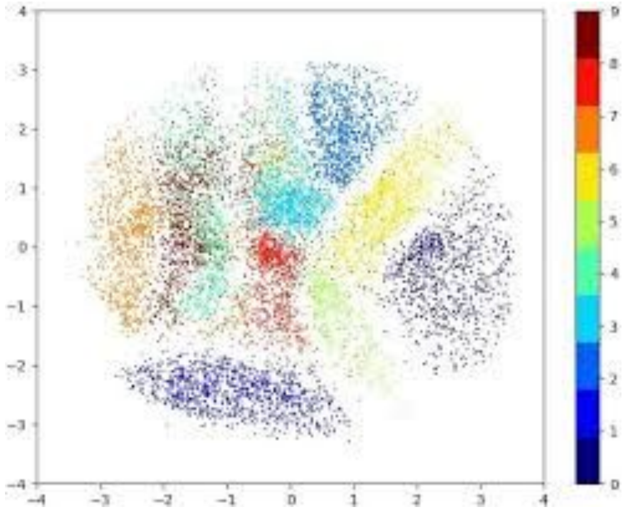
- Open space modelling: it is able to provide explicit probability estimation and visualization of unknown classes
- There are no obvious performance improvement over the **natural image** setting



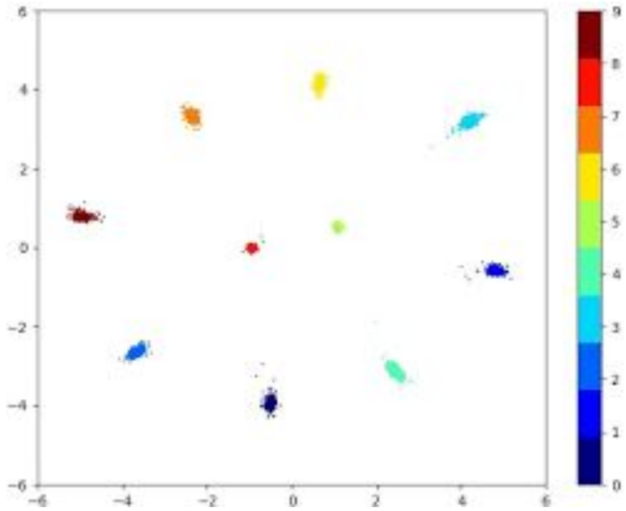
Ge, Z., Demyanov, S., Chen, Z., & Garnavi, R. (2017). Generative OpenMax for multi-class open set classification. In *British Machine Vision Conference 2017*. British Machine Vision Association and Society for Pattern Recognition.

Conditional Gaussian Distribution Learning (CGDL)

- Can classify known samples by forcing different latent features to approximate different Gaussian models



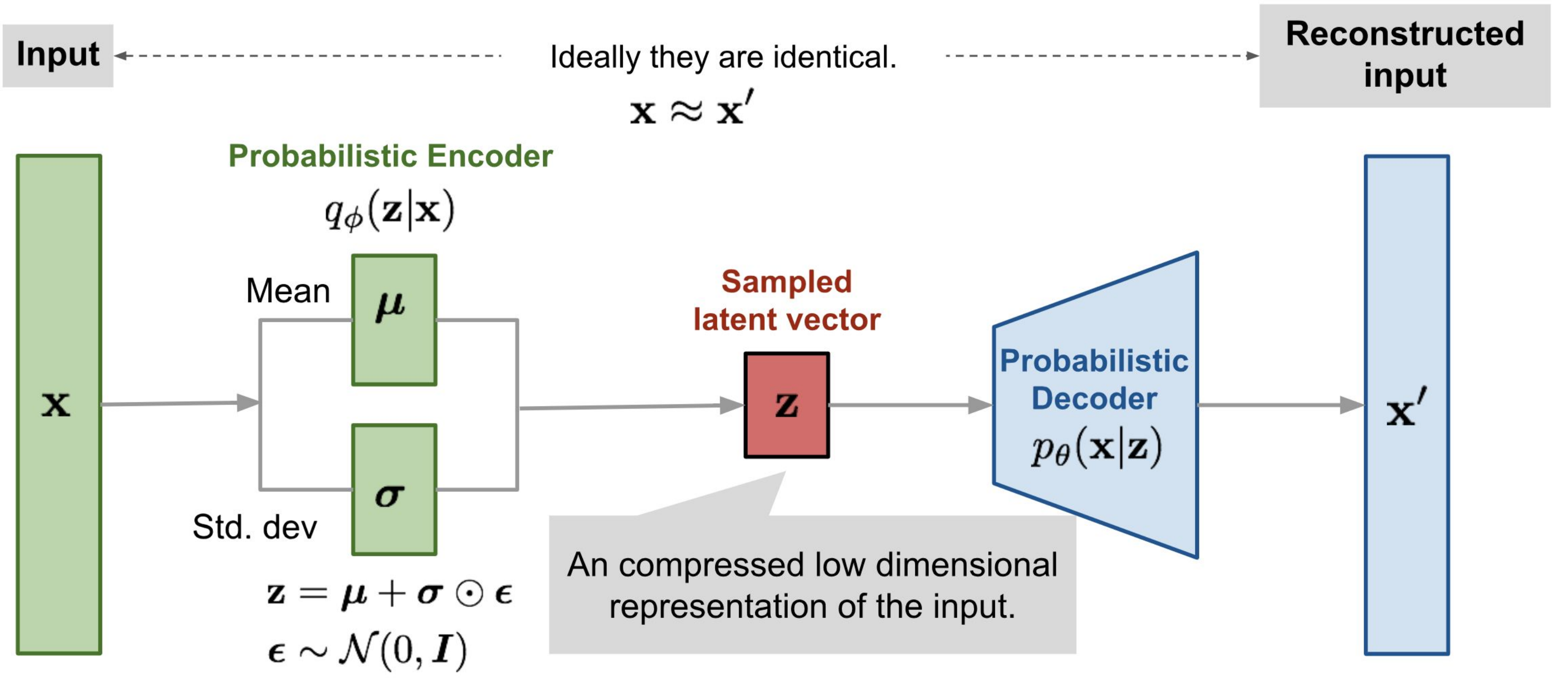
VAE



CGDL

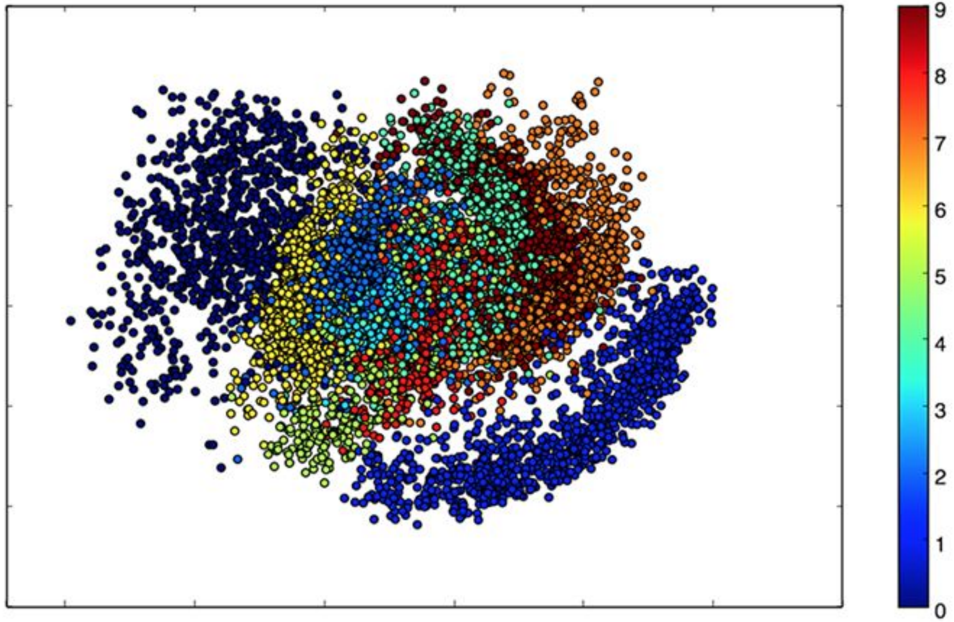
Sun, X., Yang, Z., Zhang, C., Ling, K. V., & Peng, G. (2020). Conditional Gaussian Distribution Learning for Open Set Recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 13480-13489).

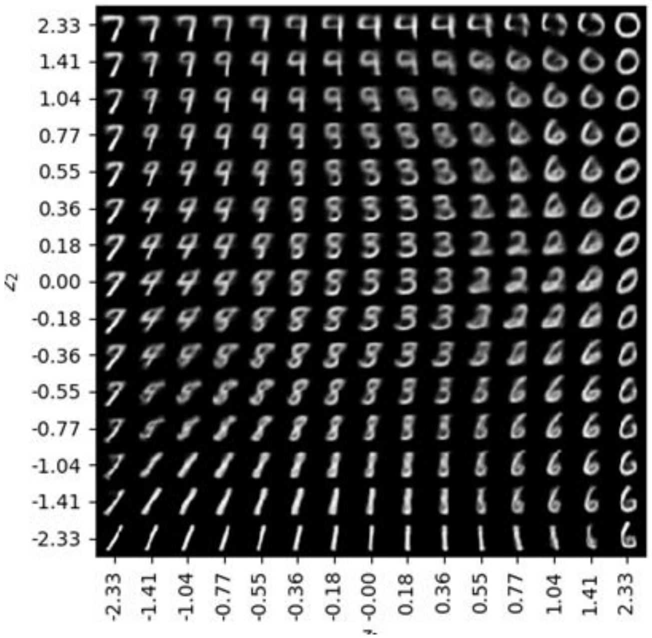
Variational Auto-Encoders



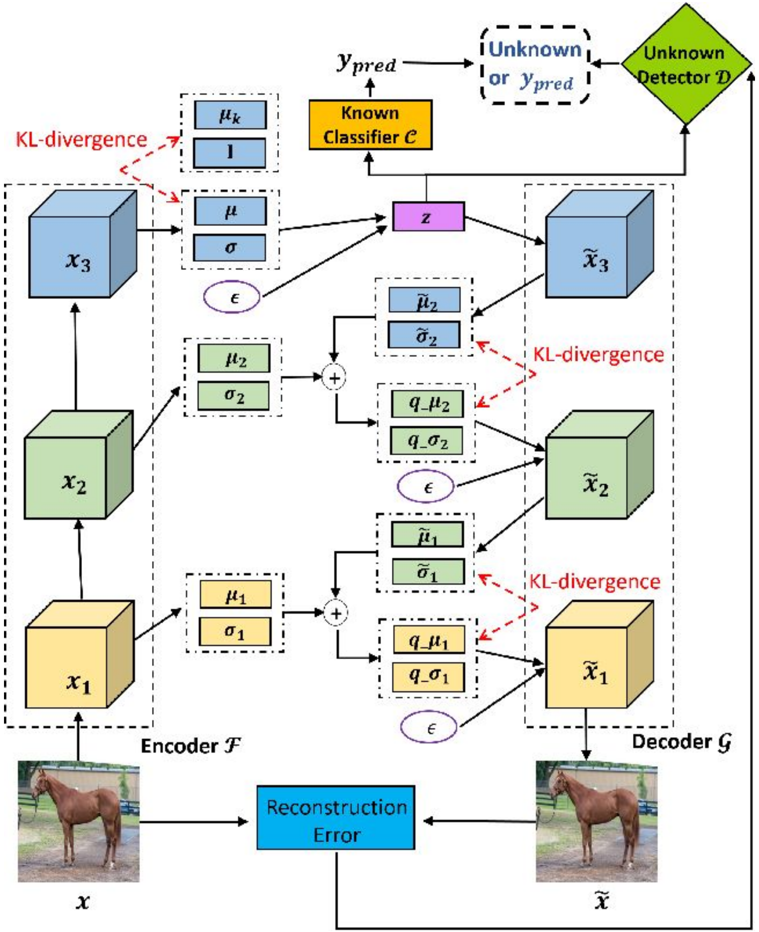
$$\boxed{\mathcal{L}(\theta, \phi) = -\mathbb{E}_{q_\phi(\mathbf{z}|\boldsymbol{x})}\left[\log p_\theta(\boldsymbol{x}|\mathbf{z})\right] + KL\big(q_\phi(\mathbf{z}|\boldsymbol{x})||p_\theta(\mathbf{z})\big)}$$

Variational Auto-Encoders





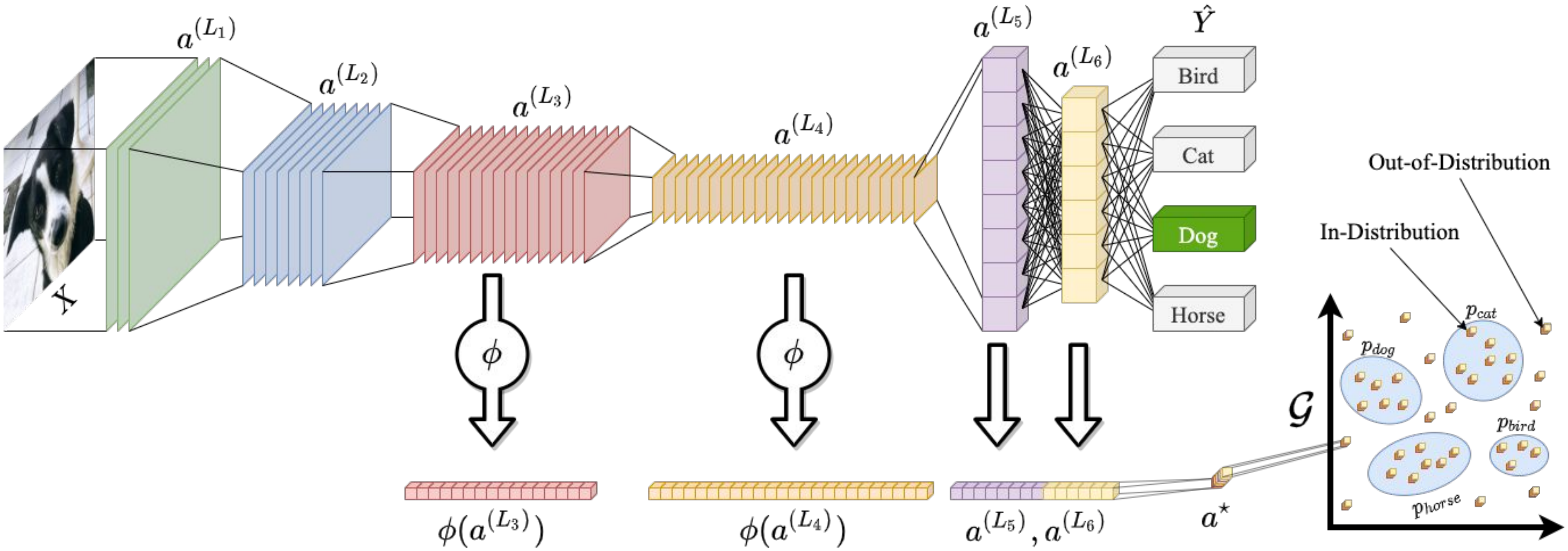
Conditional Gaussian Distribution Learning (CGDL)



Method	MNIST	SVHN	CIFAR10	CIFAR+10
Softmax	0.978 ± 0.002	0.886 ± 0.006	0.677 ± 0.032	$0.816 \pm$ -
Openmax [4]	0.981 ± 0.002	0.894 ± 0.008	0.695 ± 0.032	$0.817 \pm$ -
G-Openmax [8]	0.984 ± 0.001	0.896 ± 0.006	0.675 ± 0.035	$0.827 \pm$ -
OSRCI [22]	0.988 ± 0.001	0.910 ± 0.006	0.699 ± 0.029	$0.838 \pm$ -
C2AE [24]	0.989 ± 0.002	0.922 ± 0.009	0.895 ± 0.008	0.955 ± 0.006
ours: CGDL	$\mathbf{0.994 \pm 0.002}$	$\mathbf{0.935 \pm 0.003}$	$\mathbf{0.903 \pm 0.009}$	$\mathbf{0.959 \pm 0.006}$

Sun, X., Yang, Z., Zhang, C., Ling, K. V., & Peng, G. (2020). Conditional Gaussian Distribution Learning for Open Set Recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 13480-13489).

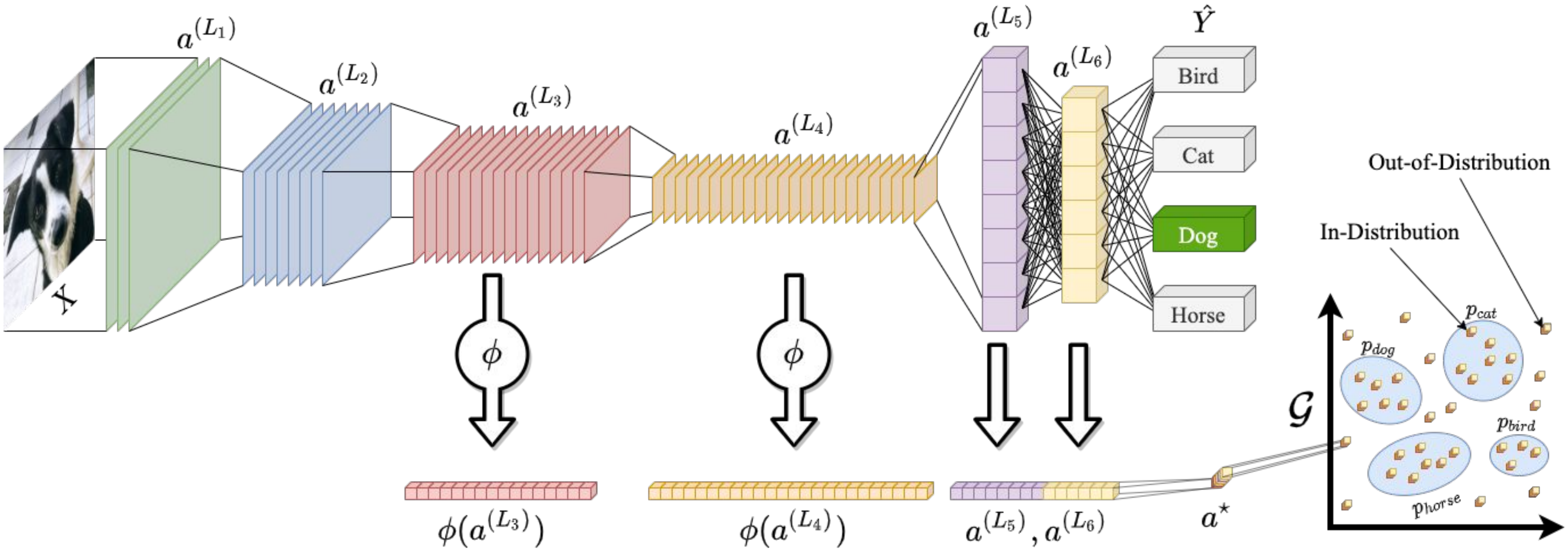
Generative Models for Open Set recognition (GeMOS)



- CNN extracts feature-level information from KKCs
- Generative models use these features to assign a score samples and identify UUCs

Vendramini, M., Oliveira, H., Machado, A., & dos Santos, J. A. (2021). Opening Deep Neural Networks with Generative Models. *ICIP.*

Generative Models for Open Set recognition (GeMOS)



- Closed set image recognition is performed by the CNN
- OSR is introduced by thresholding on likelihood scores generated by G

Vendramini, M., Oliveira, H., Machado, A., & dos Santos, J. A. (2021). Opening Deep Neural Networks with Generative Models. *ICIP.*

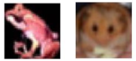
Results - CIFAR10 - D^{in}

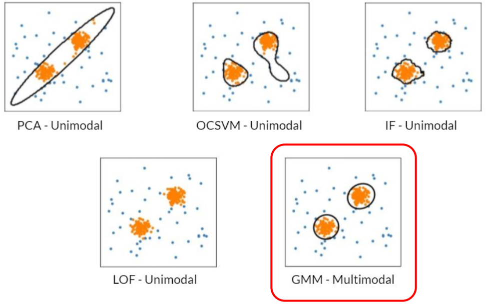
D^{out}	CF100	TIN (crop)	TIN (resize)	LSUN (crop)	LSUN (Resize)
CROSR	-	0.72	0.74	0.72	0.75
C2AE	-	0.84	0.83	0.78	0.80
CGDL	-	0.84	0.83	0.81	0.81
GeMOS	**0.80**	**0.92**	**0.92**	**0.90**	**0.93**

F1-Score

D^{out}	CF100	TIN (crop)	TIN (resize)	LSUN (crop)	LSUN (Resize)
ODIN	**0.90**	0.94	0.92	0.96	0.95
OODNN	-	0.96	0.95	**0.97**	0.96
MSP	0.88	-	-	-	-
GeMOS	0.88	**0.98**	**0.99**	**0.97**	**0.99**

AUC

- CIFAR10 (D^{in}): 
  - CIFAR100 (D^{out}): 
  - Tiny-Imagenet (D^{out}): 
  - LSUN (D^{out}): 



- **Best architecture models:** WRN-28-10 + GMM8

Generative Models for Open Set recognition (GeMOS)

- Advantage:
  - Generative models can be attached to any pre-trained DNN
  - Generative models do not require GPU for training

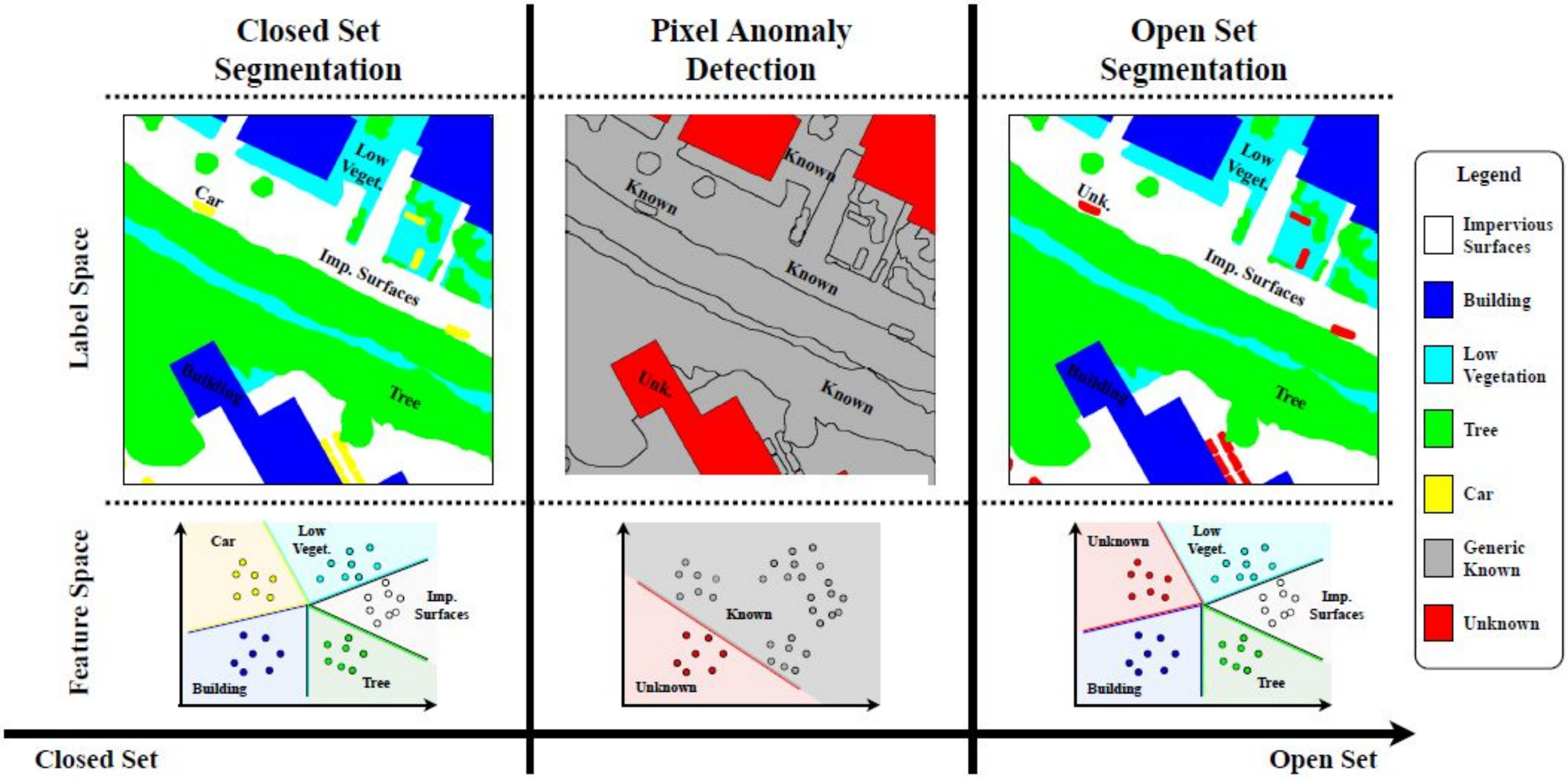
- Drawback:
  - GeMOS has a high reliance on the performance of the DNN

Experimental Results - Future

Deep-based Methods

- Would hybrid models be possible (generative and discriminative)?
- Unify clustering and classification models since the former better defines the boundaries between classes while the latter has better discriminative capacity.
- How to generate valid UUC data?
- Interpretation of rejected data.
- Use of secondary information to identify, discriminate or define unknown classes.

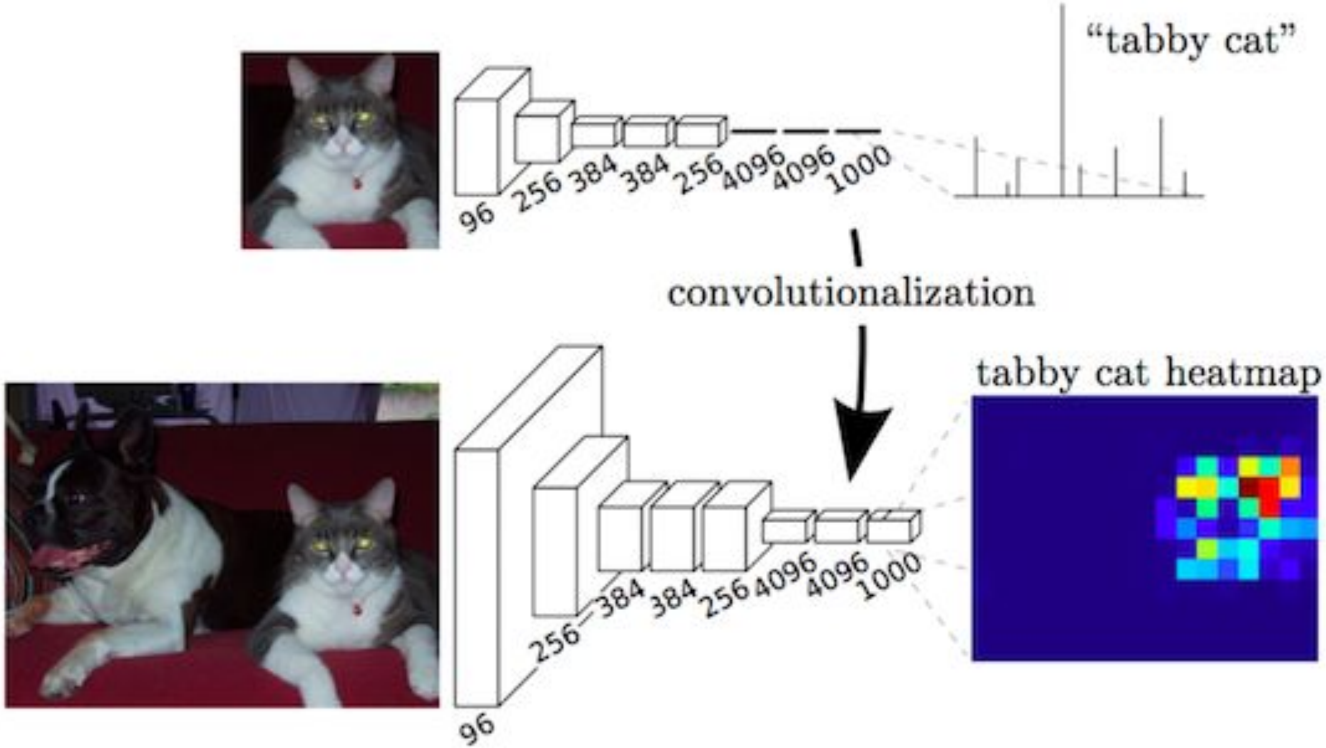
Segmentation Methods



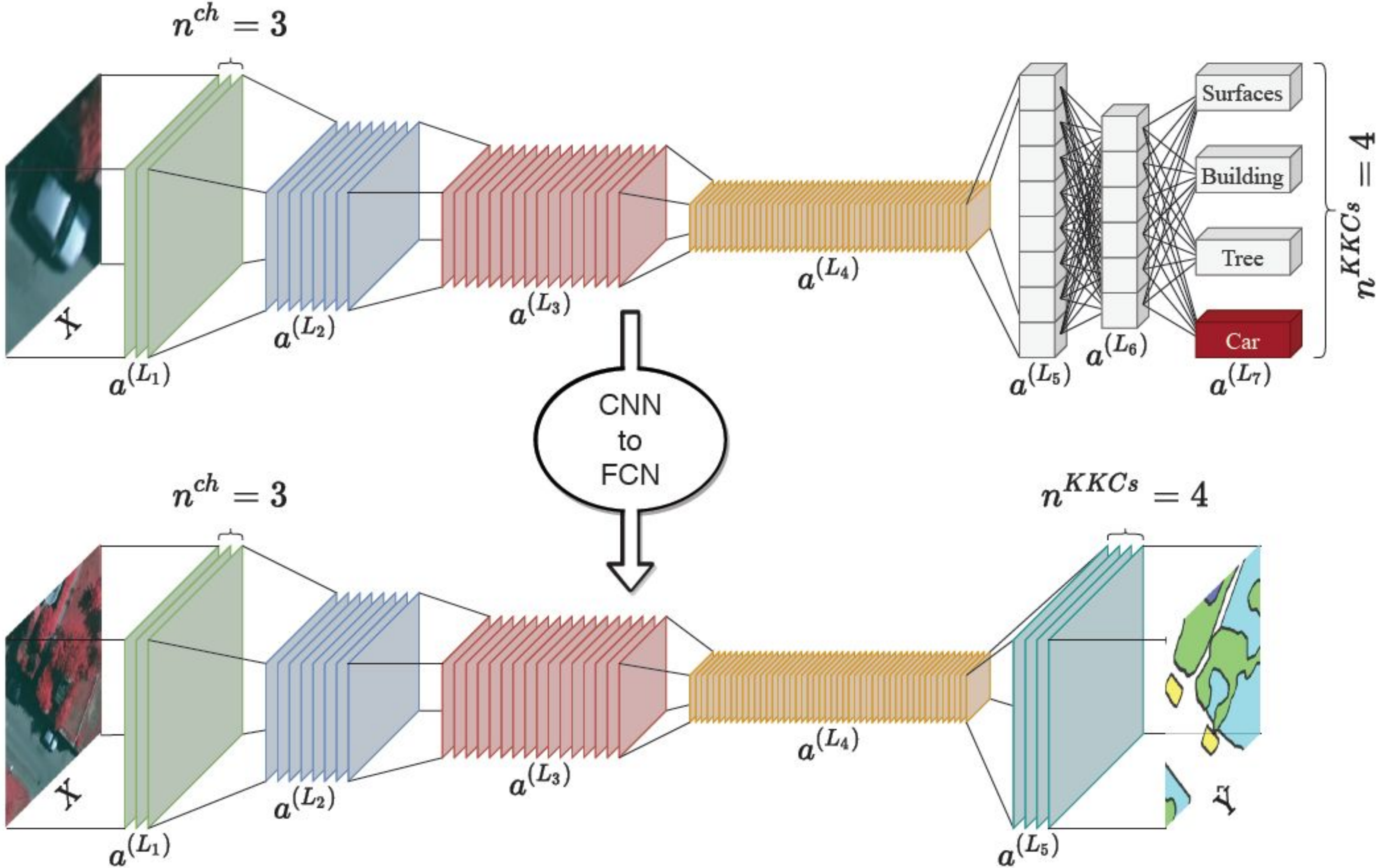
Closed Set Segmentation
Pixel Anomaly Detection
Open Set Segmentation
Label Space
Feature Space
Low Veget.
Car
Imp. Surfaces
Building
Tree
Known
Unk.
Low Veget.
Car
Imp. Surfaces
Building
Tree
Known
Unknown
Unknown
Legend
Impervious Surfaces
Building
Low Vegetation
Tree
Car
Generic Known
Unknown
Closed Set
Open Set

Semantic Segmentation

Overview

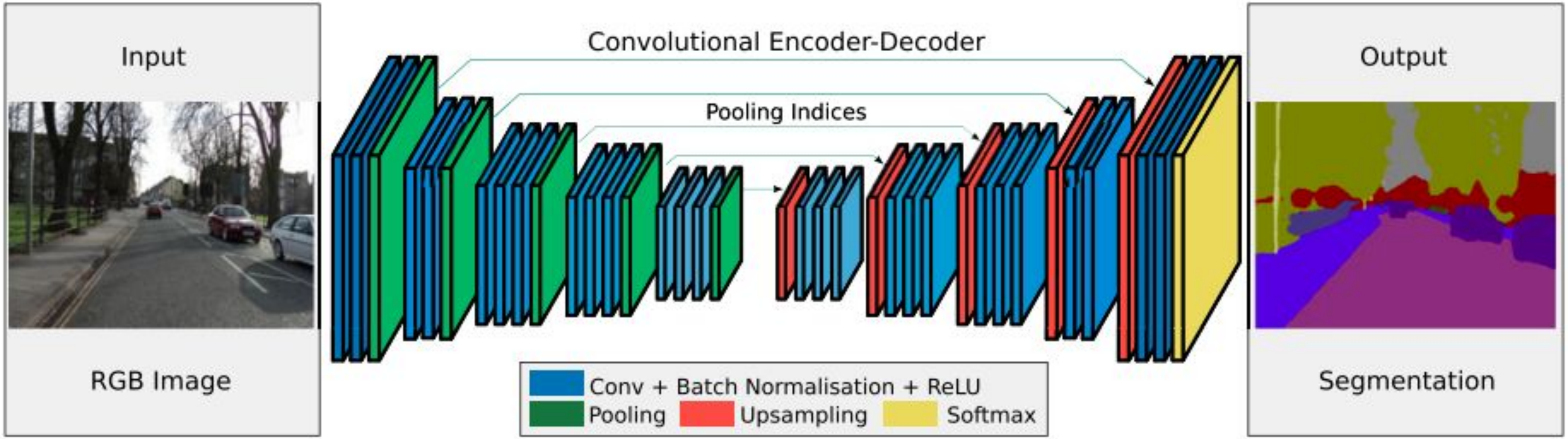


Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).



$n^{ch} = 3$
X
$a^{(L_1)}$
$a^{(L_2)}$
$a^{(L_3)}$
$a^{(L_4)}$
$a^{(L_5)}$
$a^{(L_6)}$
$a^{(L_7)}$
Surfaces
Building
Tree
Car
$n^{KKCs} = 4$
CNN to FCN
$n^{ch} = 3$
X
$a^{(L_1)}$
$a^{(L_2)}$
$a^{(L_3)}$
$a^{(L_4)}$
$a^{(L_5)}$
$n^{KKCs} = 4$
Y

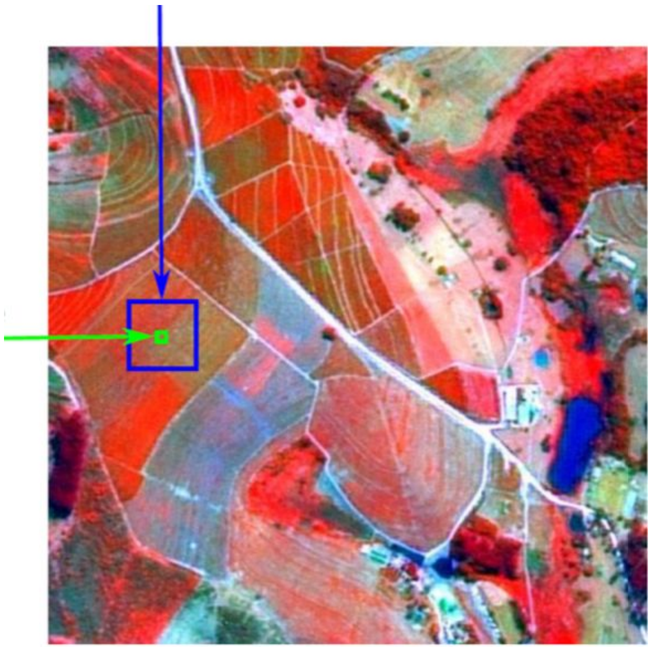
Semantic Segmentation



Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, *39*(12), 2481-2495.

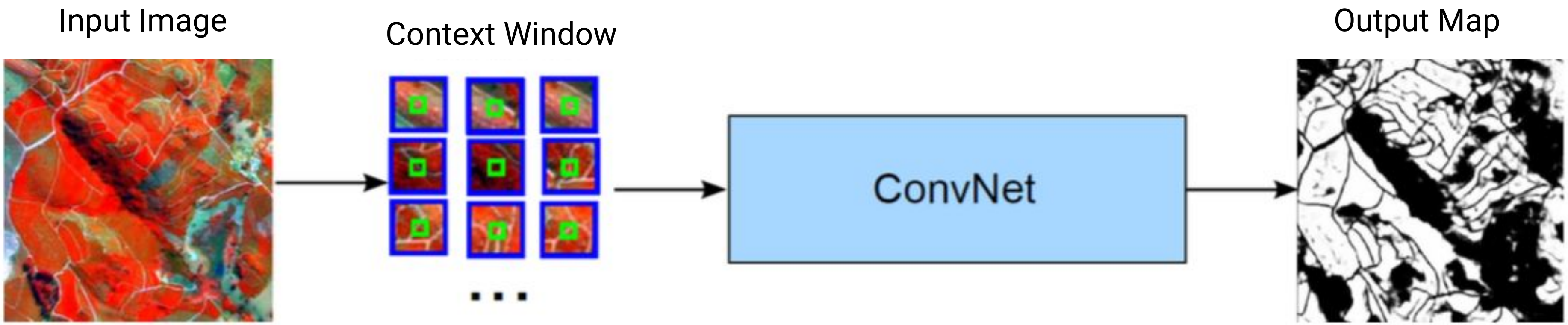
Semantic Segmentation

Pixel-Wise Classification



Semantic Segmentation

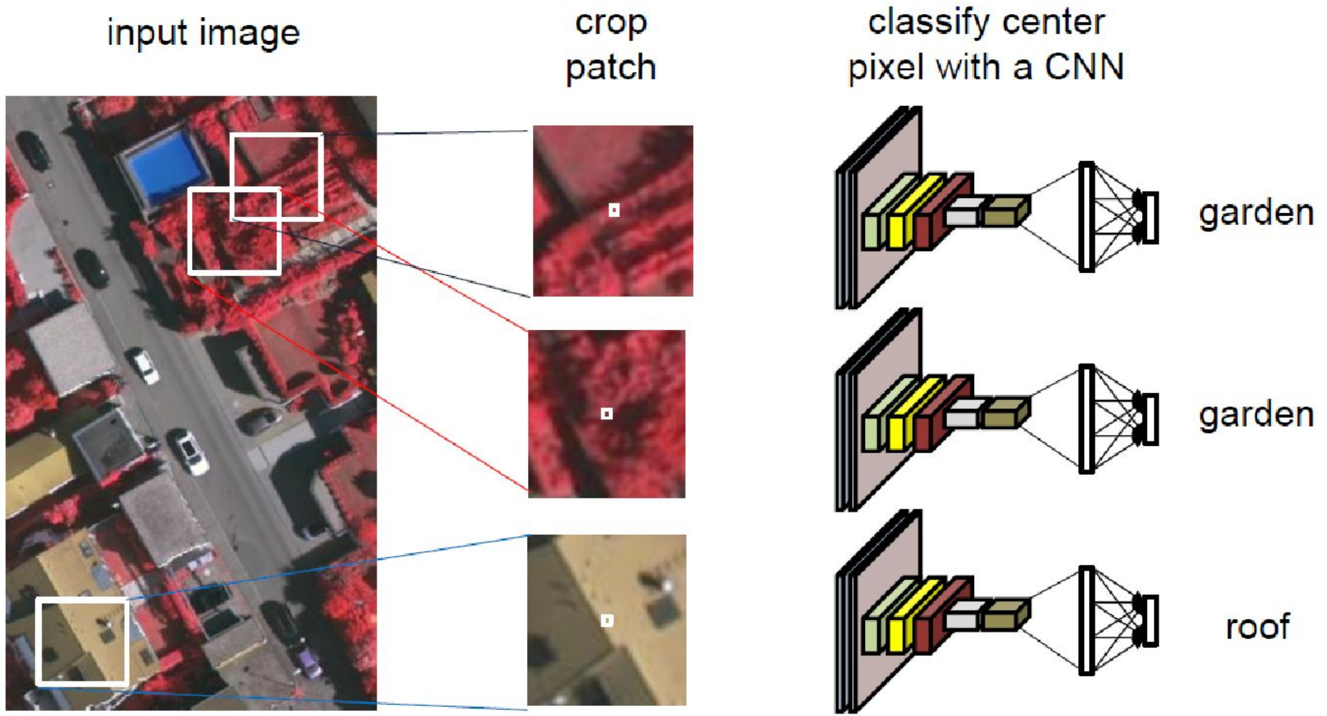
Pixel-Wise Classification



Nogueira, K., Dalla Mura, M., Chanussot, J., Schwartz, W. R., & dos Santos, J. A. (2016, December). Learning to semantically segment high-resolution remote sensing images. In *2016 23rd International Conference on Pattern Recognition (ICPR)* (pp. 3566-3571). IEEE.

Semantic Segmentation

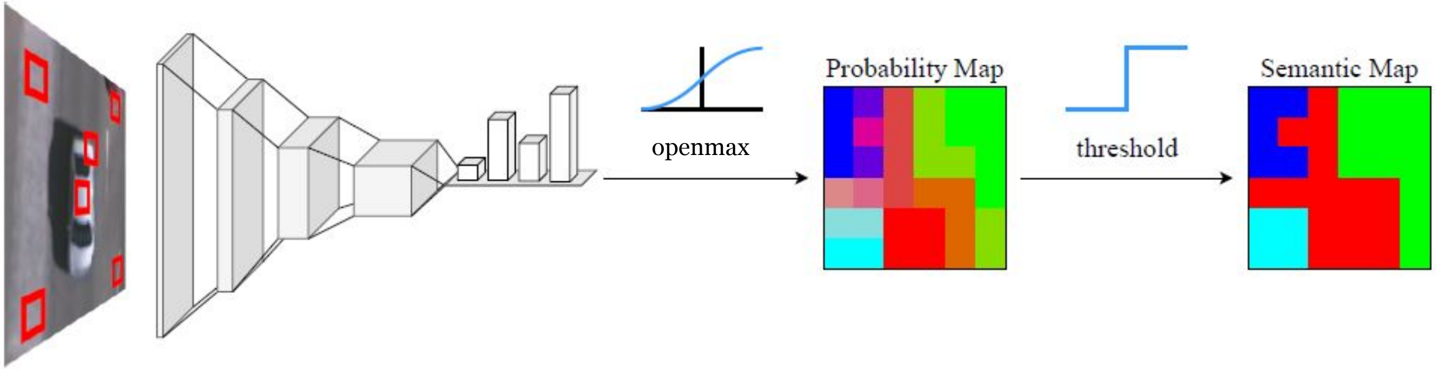
Pixel-Wise Classification



Many overlaps. Inefficient!

Pixel-Wise Open-Set Semantic Segmentation

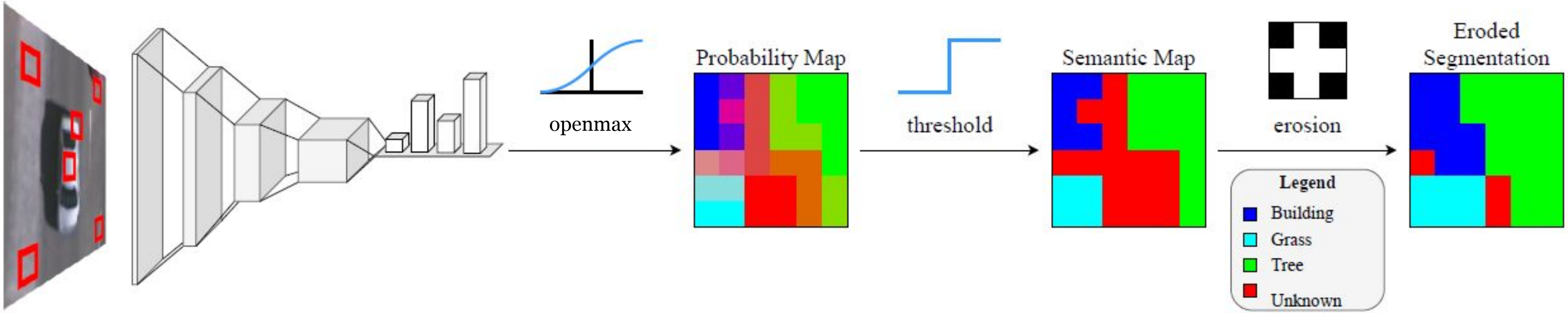
OpenPixel



da Silva, C. C., Nogueira, K., Oliveira, H. N., & dos Santos, J. A. (2020, March). Towards open-set semantic segmentation of aerial images. In *2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS)* (pp. 16-21).

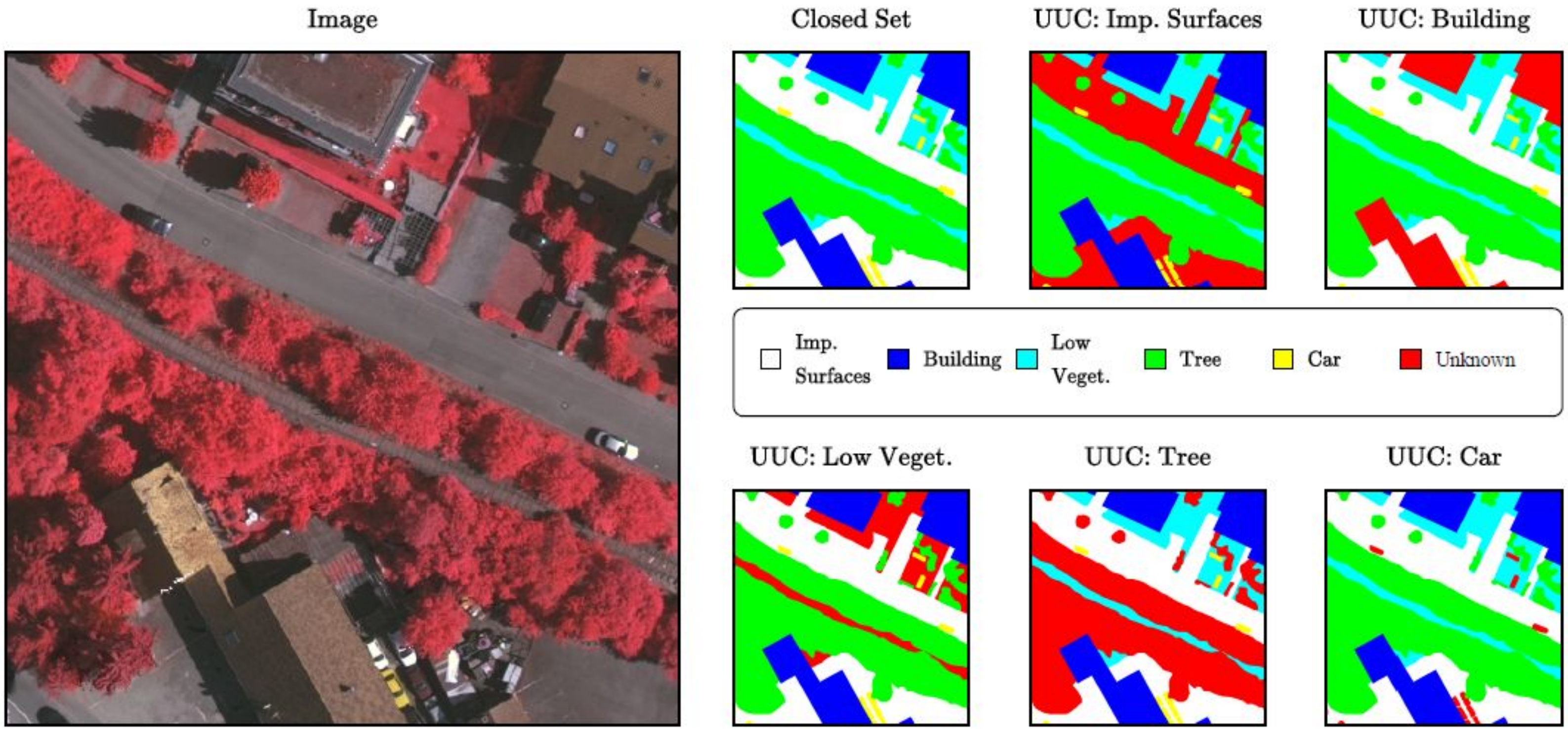
Pixel-Wise Open-Set Semantic Segmentation

Morph-OpenPixel



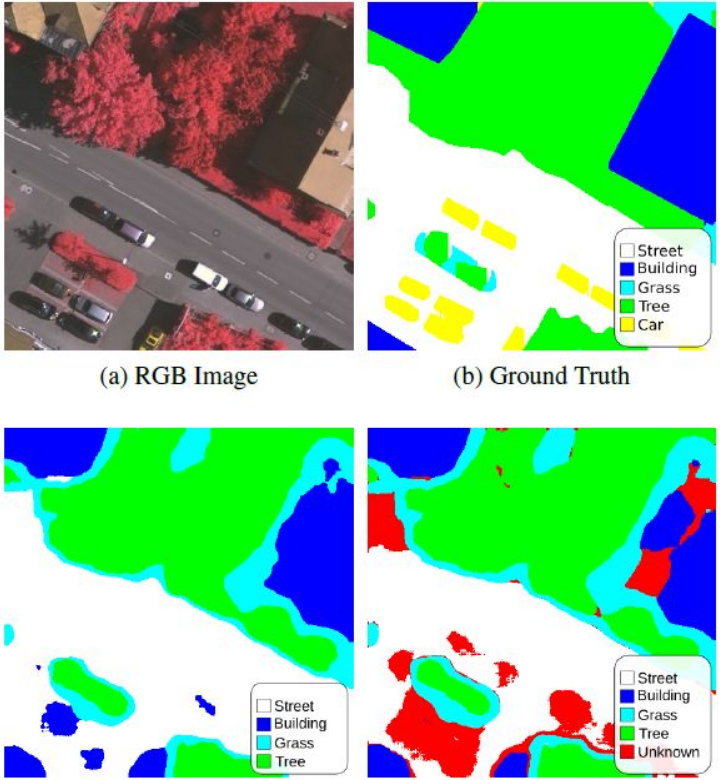
da Silva, C. C., Nogueira, K., Oliveira, H. N., & dos Santos, J. A. (2020, March). Towards open-set semantic segmentation of aerial images. In *2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS)* (pp. 16-21).

Dataset: Vaihingen



http://www2.isprs.org/commissions/comm3/wg4/2d-sem-label-vaihingen.html

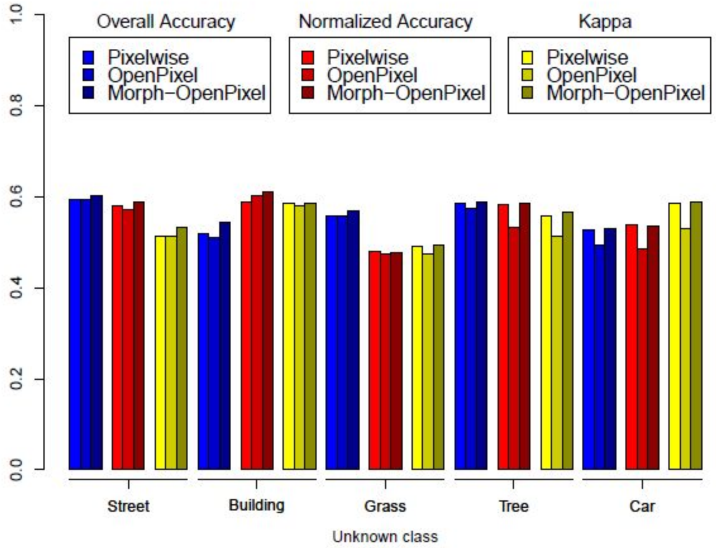
Experimental Results



(a) RGB Image (b) Ground Truth

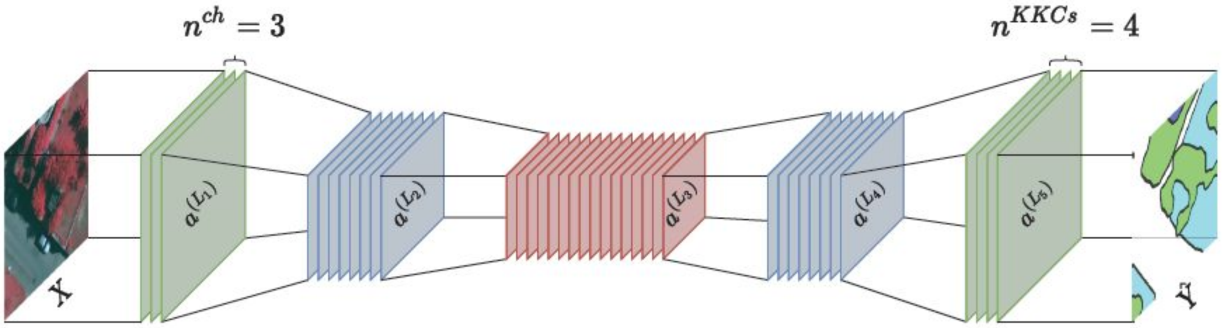
(c) Pixelwise Closed Set Prediction (d) Morph-OpenPixel Prediction

Network	Overall Accuracy	Normalized Accuracy	Kappa
Pixelwise (closed)	55.84%	53.98%	0.5585
OpenPixel	55.78%	53.15%	0.5106
Morph-OpenPixel	**57.51%**	**54.23%**	**0.5602**



Fully Convolutional Open-Set Semantic Segmentation

Open-FCN



Let $\{X, Y\}$ be a paired set of image pixels and semantic labels from a dataset containing C KKCs. A deep model $\mathcal{M}$ can be trained in a stochastic manner by feeding samples X_i to a gradient descent optimizer as Adam [18] with a loss function as Cross Entropy, given by:

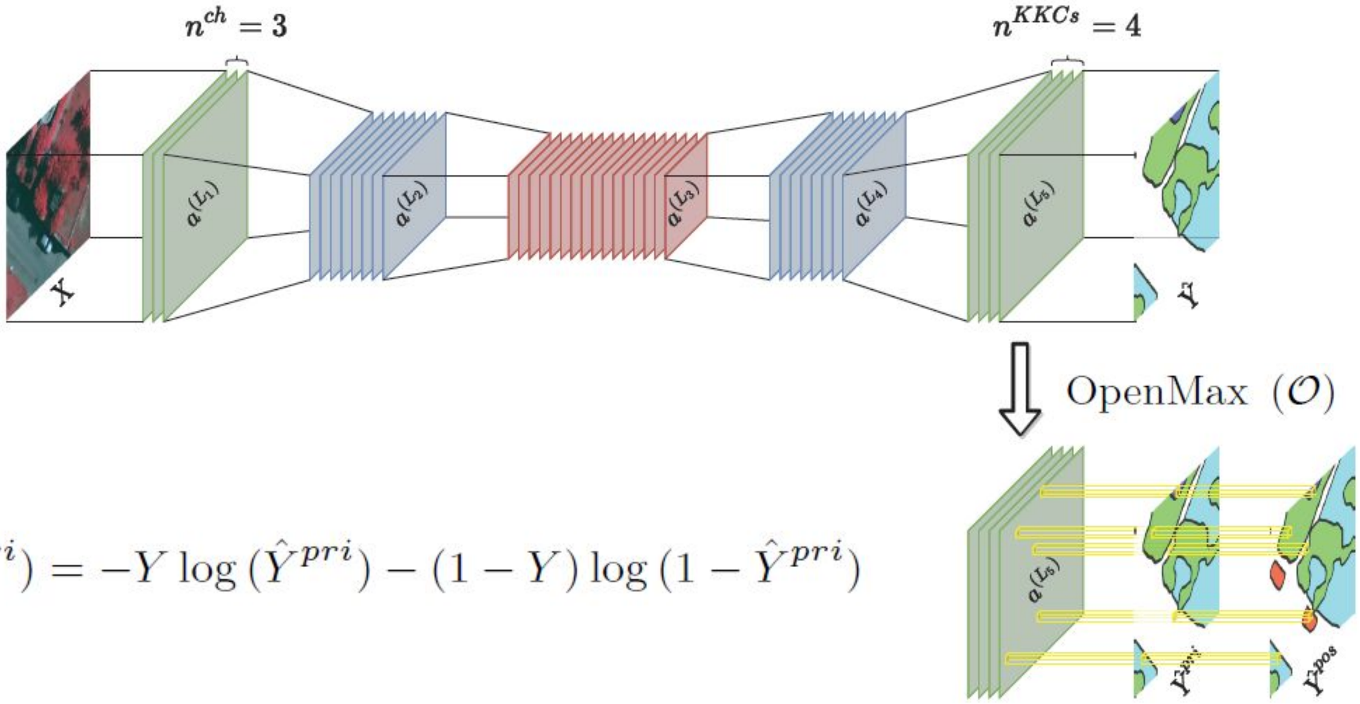
$$\mathcal{L}_{CE}(Y, \hat{Y}^{pri}) = -Y \log(\hat{Y}^{pri}) - (1 - Y)\log(1 - \hat{Y}^{pri})$$

Training (only Known Known Classes - KKCs)

Oliveira, H., Silva, C., Machado, G. L., Nogueira, K., & Santos, J. A. dos (2020). Fully Convolutional Open Set Segmentation. *arXiv preprint arXiv:2006.14673*.

Fully Convolutional Open-Set Semantic Segmentation

Open-FCN



$$\mathcal{L}_{CE}(Y, \hat{Y}^{pri}) = -Y \log(\hat{Y}^{pri}) - (1 - Y) \log(1 - \hat{Y}^{pri})$$

Oliveira, H., Silva, C., Machado, G. L., Nogueira, K., & Santos, J. A. dos (2020). Fully Convolutional Open Set Segmentation. *arXiv preprint arXiv:2006.14673*.

OpenMax Computation

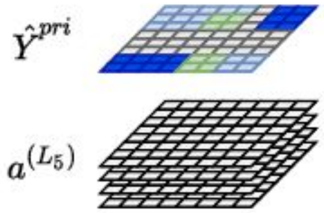
- It adjusts softmax probabilities
- It is performed on the validation phase:

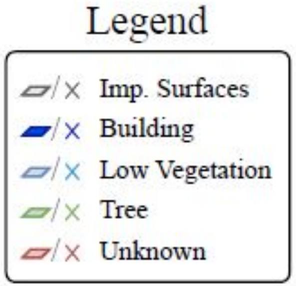
Each $\text{KKC } c_k, k \in \{0, 1, ..., C-1\}$ yields one Weibull distribution $\mathcal{W}_k$.

$\mathcal{W}_k$ is t to the deviations from the mean μ_k of $a^{(L_5)}$

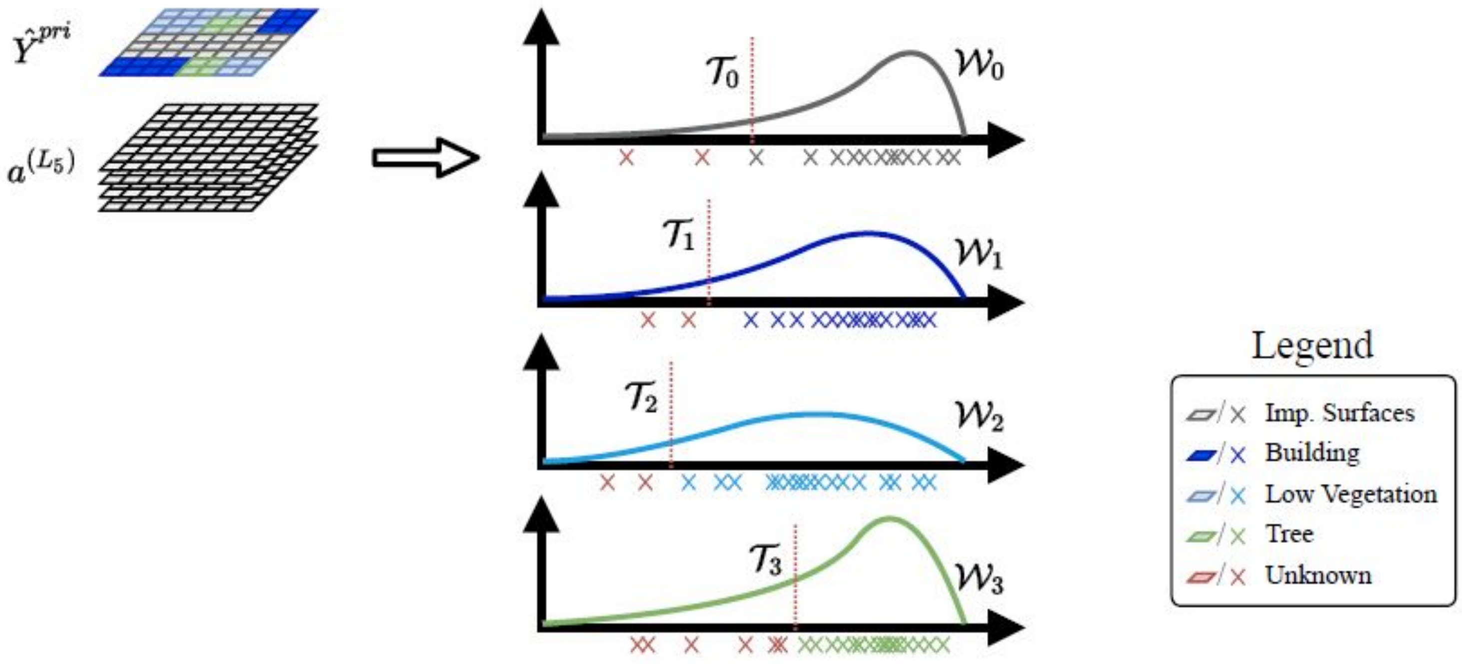
In order to identify Out-of-Distribution (OOD) samples, quantiles from the Cumulative Distribution Function (CDF) for $\mathcal{W}_{||}$ is computed

KKC and UUC Prediction





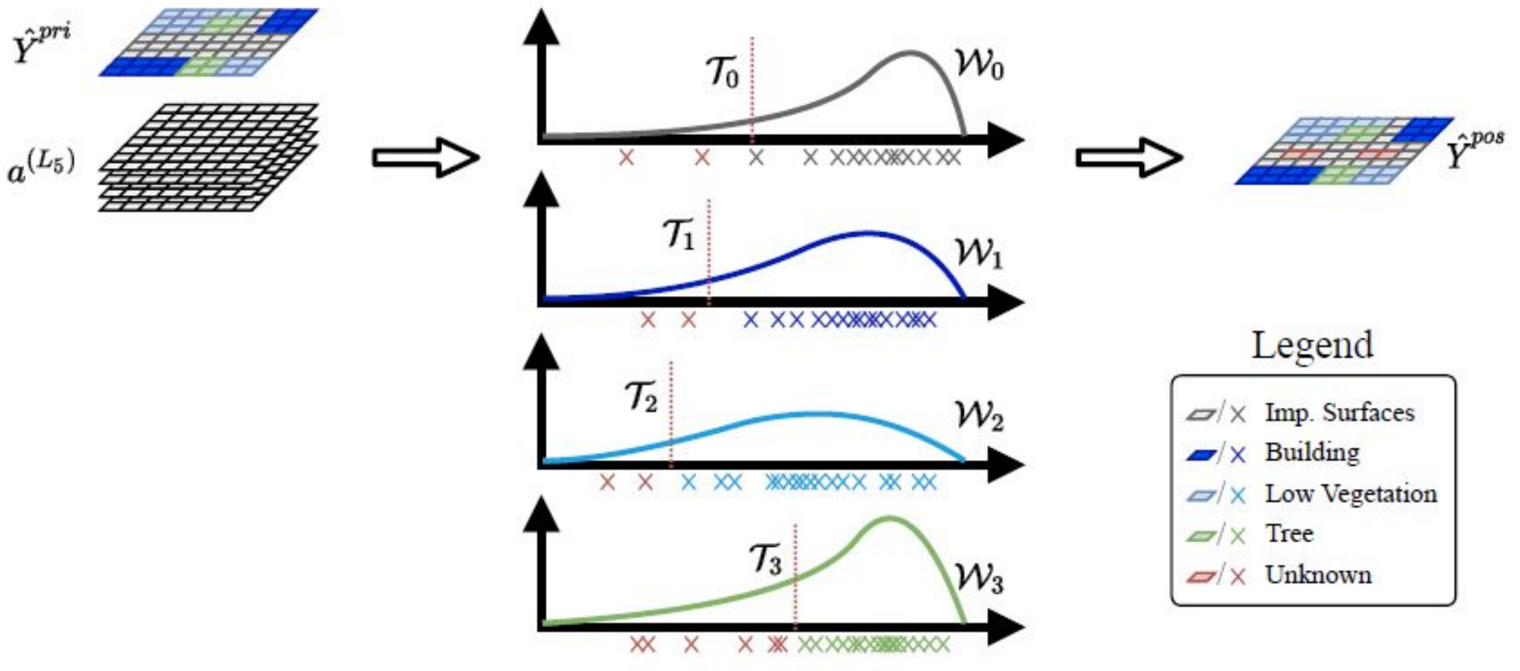
KKC and UUC Prediction



KKC and UUC Prediction

The posterior OpenMax prediction for a specific pixel Xi is given by:

$$\hat{Y}_i^{pos} = \begin{cases} c_k, \text{ if } max(a_i^{(l)}) \geq \mathcal{T}_k \\ c_{unk}, \text{ if } max(a_i^{(l)}) < \mathcal{T}_k, \end{cases}$$



KKC and UUC Prediction

Limitations

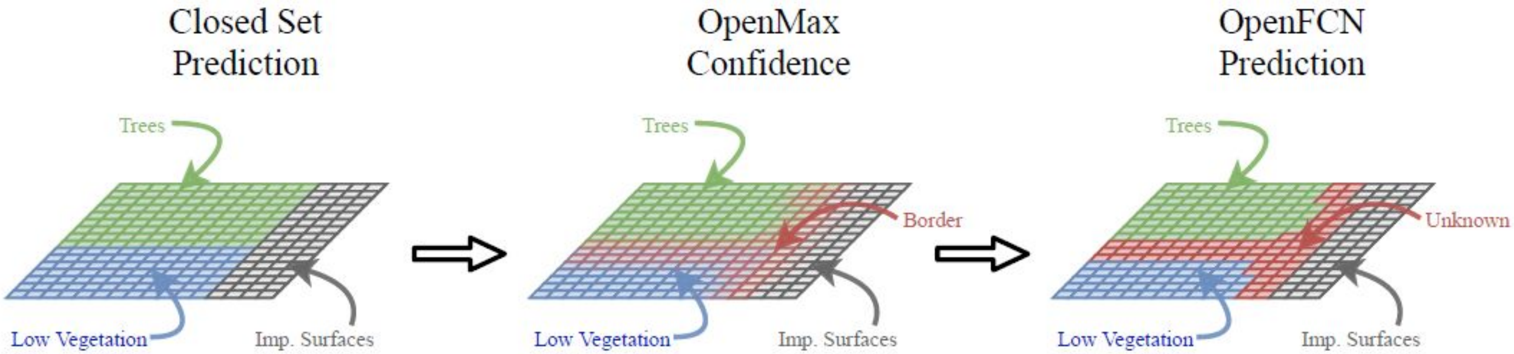
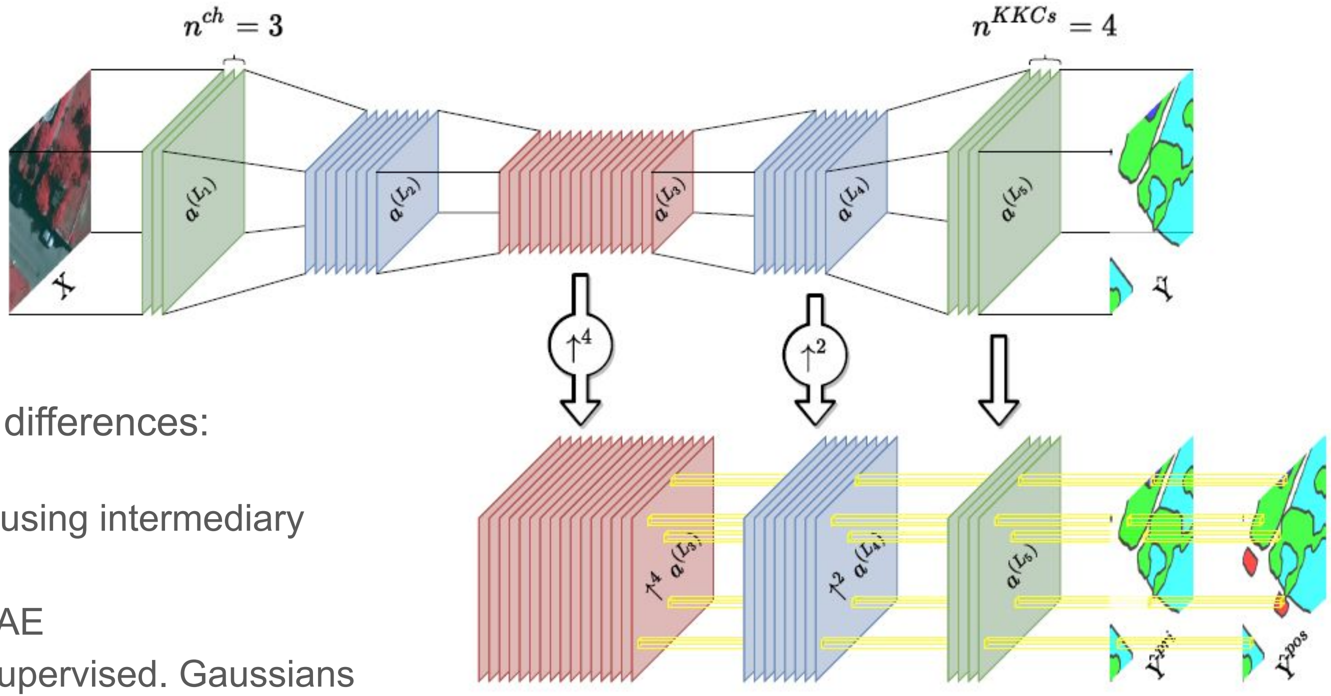


Fig. 5: Depiction of OpenFCN's prediction confidence degradation on object boundaries due to the use of information close to the label space. SoftMax and OpenMax probabilities on dense labeling tasks are naturally lower on boundary regions between objects from distinct classes.

Open Set Scoring with Principal Components

OpenPCS



It extends CGDL. Main differences:

1. Fit gaussian priors using intermediary activations
2. PCA instead of a VAE
3. Training is purely supervised. Gaussians are fitted in the prediction phase

Oliveira, H., Silva, C., Machado, G. L., Nogueira, K., & Santos, J. A. dos (2020). Fully Convolutional Open Set Segmentation. *arXiv preprint arXiv:2006.14673*.

Why employing earlier layers?

The deeper a certain layer l is placed in a DNN, the closer to the label space the activation features $a(l)$.

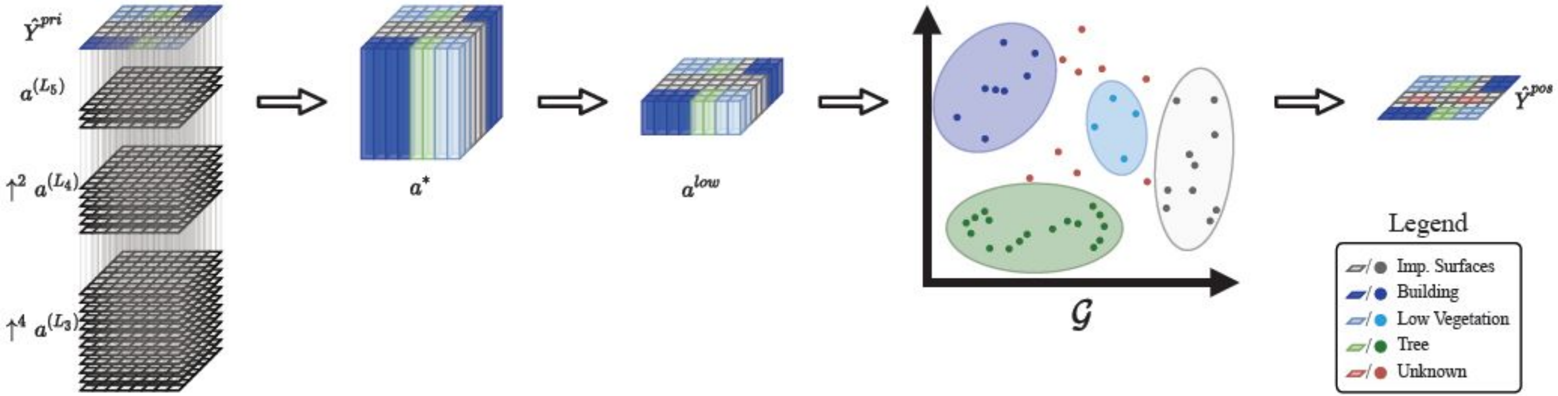
Shwartz and Tishby argue that any supervised DNN can be seen as a Markov chain of sequential tensorial representations that gradually morph the information processed by the network from the input space (in the input layer) to the label space (in the output layer).

Thus, by using only the last layer's activations to t Weibull distributions to each KKC, OpenFCNs limit themselves to work with information close to the label space.

Shwartz-Ziv, R., & Tishby, N. (2017). Opening the black box of deep neural networks via information. *arXiv preprint arXiv:1703.00810*.

Principal Component Analysis (PCA)

1. Faster inference time during testing
   - PCA implementations can be highly parallelized via vectorial operations and low-dimensional gaussian likelihood scoring can be computed in a fast manner
2. Only the most important activation channels are used to compute a scoring function to detect out of data samples



Experimental Results

AUC - Vaihingen Dataset

Backbone	Methods	UUC: Imp. Surf.	UUC: Building	UUC: Low Veg.	UUC: Tree	UUC: Car
DenseNet	SoftMax^{T}	.78 ± .03	.63 ± .05	**.73 ± .05**	.67 ± .06	.58 ± .06
	OpenFCN	.81 ± .03	.64 ± .05	.72 ± .05	.66 ± .06	.58 ± .05
	OpenPCS	**.84 ± .03**	**.94 ± .01**†	.70 ± .08	**.81 ± .07**†	**.65 ± .06**
WRN	SoftMax^{T}	.80 ± .03	.64 ± .04	.75 ± .05	**.63 ± .06**	.58 ± .05
	OpenFCN	.83 ± .03	.67 ± .03	.74 ± .05	.62 ± .06	.58 ± .05
	OpenPCS	**.87 ± .03**	**.94 ± .02**†	**.77 ± .04**	.63 ± .09	**.87 ± .03**†
ResNeXt	SoftMax^{T}	.66 ± .06	.55 ± .07	**.71 ± .07**	.66 ± .04	.65 ± .04
	OpenFCN	.67 ± .06	.59 ± .05	.71 ± .07	.64 ± .03	.63 ± .04
	OpenPCS	**.88 ± .03**†	**.81 ± .07**	.58 ± .08	**.69 ± .09**	**.65 ± .05**
ResNet	SoftMax^{T}	.83 ± .03	.62 ± .06	**.81 ± .05**†	.60 ± .05	.60 ± .05
	OpenFCN	**.86 ± .02**	.64 ± .05	.80 ± .05	**.61 ± .05**	.62 ± .05
	OpenPCS	.75 ± .03	**.92 ± .02**	.65 ± .07	.50 ± .07	**.63 ± .05**
VGG	SoftMax^{T}	.72 ± .02	.59 ± .04	.74 ± .04	**.66 ± .05**	.66 ± .05
	OpenFCN	.74 ± .03	.59 ± .04	**.74 ± .03**	.65 ± .05	.65 ± .04
	OpenPCS	**.82 ± .04**	**.90 ± .02**	.71 ± .06	.60 ± .07	**.81 ± .04**

Experimental Results

AUC - Potsdam Dataset

Backbone	Methods	UUC: Imp. Surf.	UUC: Building	UUC: Low Veg.	UUC: Tree	UUC: Car
DenseNet	SoftMax[7]	$.66 \pm .09$	$.49 \pm .07$	$\mathbf{.53 \pm .13}$	$\mathbf{.64 \pm .08}$†	$.76 \pm .02$
	OpenFCN	$.66 \pm .09$	$.50 \pm .06$	$.51 \pm .13$	$.62 \pm .07$	$.80 \pm .03$
	OpenPCS	$.83 \pm .06$	$.87 \pm .13$	$.28 \pm .03$	$.54 \pm .11$	$.77 \pm .04$
	OpenIPCS	$\mathbf{.84 \pm .07}$†	$\mathbf{.88 \pm .14}$†	$.41 \pm .07$	$.54 \pm .09$	$\mathbf{.91 \pm .04}$†
WRN	SoftMax[7]	$.68 \pm .06$	$.53 \pm .06$	$.60 \pm .12$	$\mathbf{.62 \pm .07}$	$.66 \pm .10$
	OpenFCN	$\mathbf{.69 \pm .05}$	$.54 \pm .05$	$.58 \pm .12$	$.60 \pm .07$	$.69 \pm .09$
	OpenPCS	$.66 \pm .08$	$\mathbf{.86 \pm .10}$	$.71 \pm .08$	$.57 \pm .10$	$\mathbf{.86 \pm .04}$
	OpenIPCS	$.64 \pm .09$	$\mathbf{.86 \pm .10}$	$\mathbf{.76 \pm .07}$†	$.55 \pm .10$	$.83 \pm .04$

VAIHINGEN

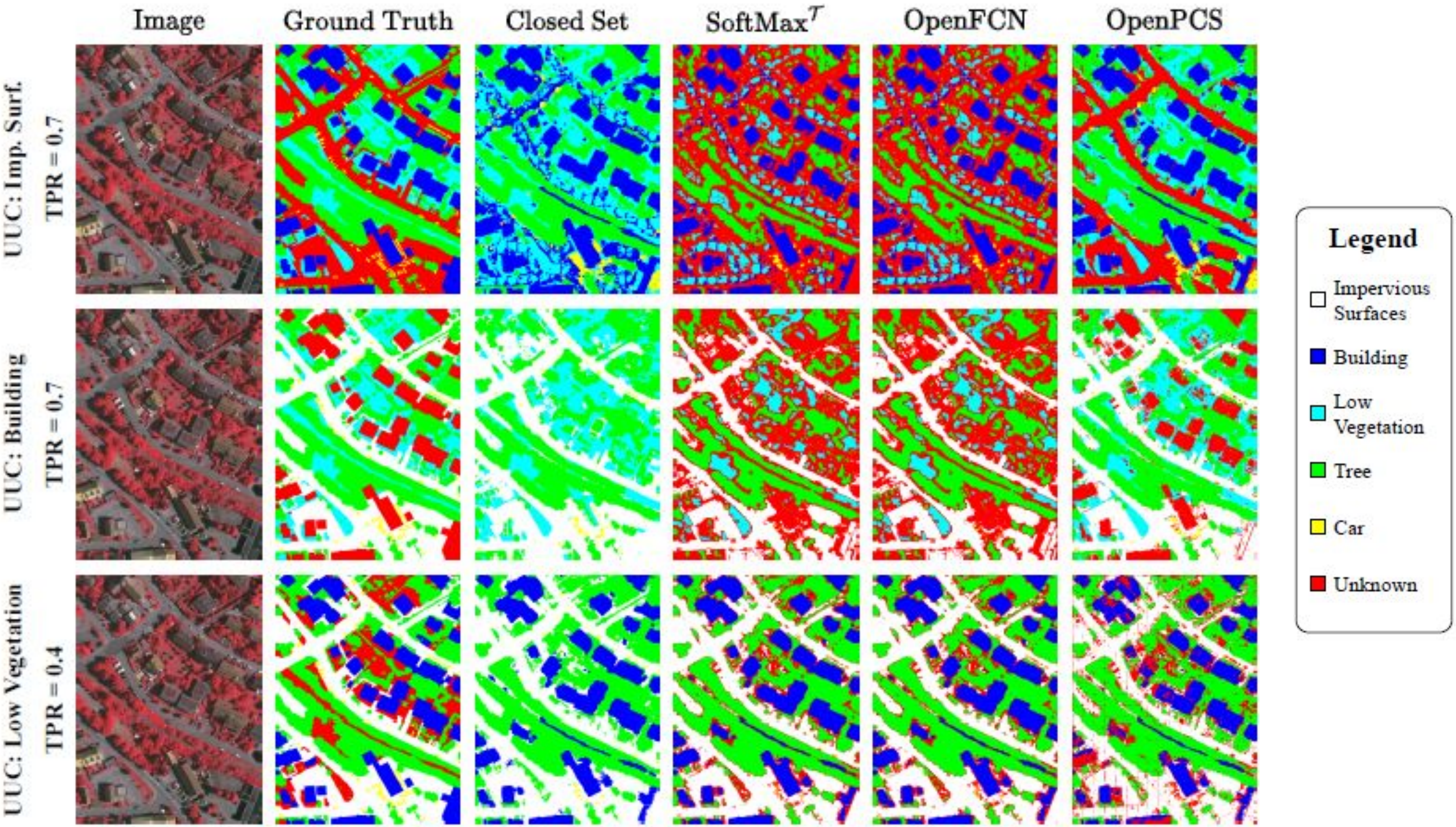


Image
Ground Truth
Closed Set
SoftMax^T
OpenFCN
OpenPCS
UUC: Imp. Surf.
TPR = 0.7
UUC: Building
TPR = 0.7
UUC: Low Vegetation
TPR = 0.4
Legend
Impervious Surfaces
Building
Low Vegetation
Tree
Car
Unknown

VAIHINGEN

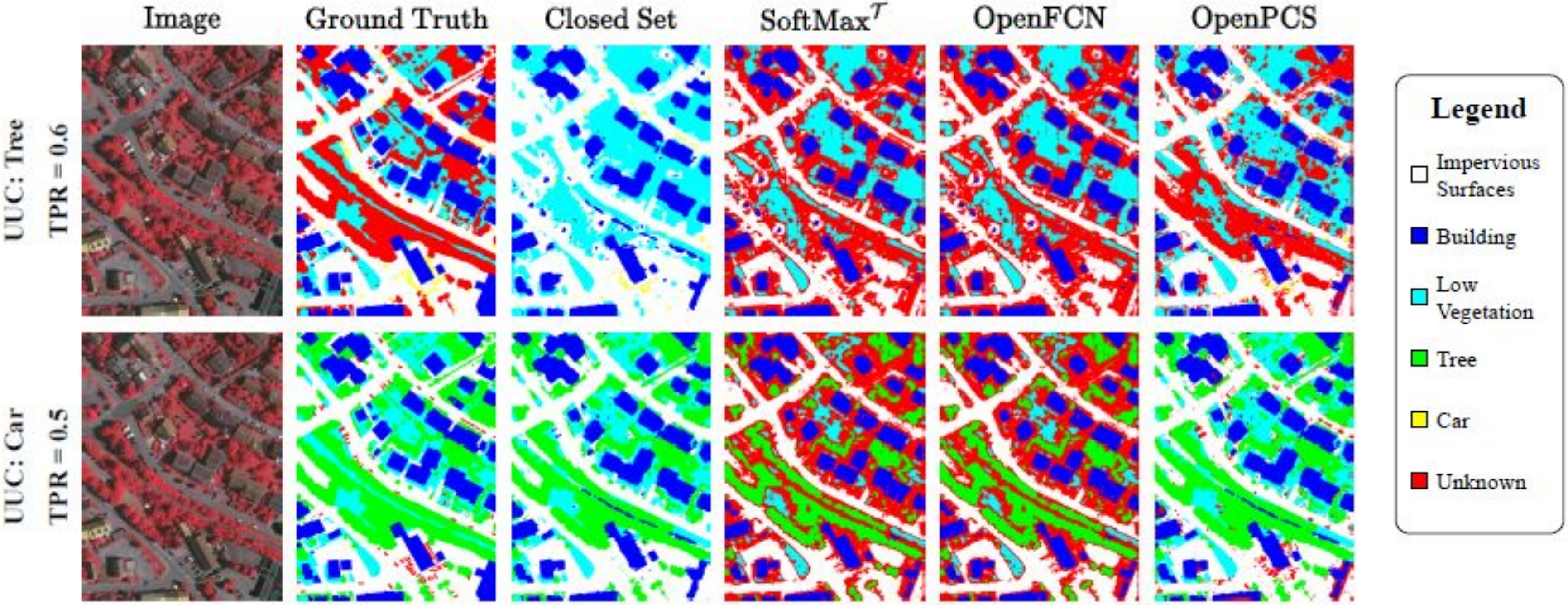


Image
Ground Truth
Closed Set
SoftMax^T
OpenFCN
OpenPCS
UUC: Tree
TPR = 0.6
UUC: Car
TPR = 0.5
Legend
Impervious Surfaces
Building
Low Vegetation
Tree
Car
Unknown

POTSDAM

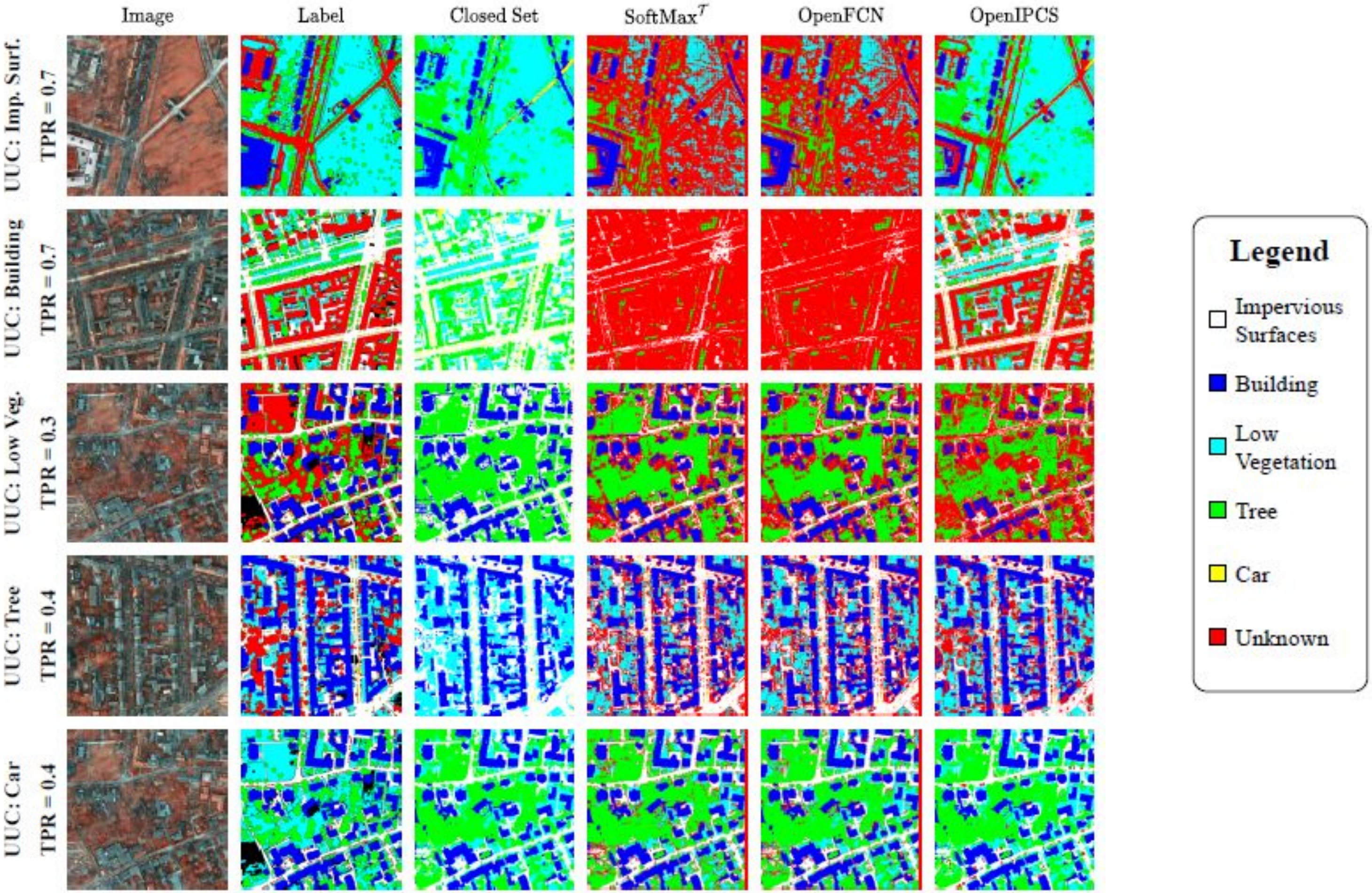
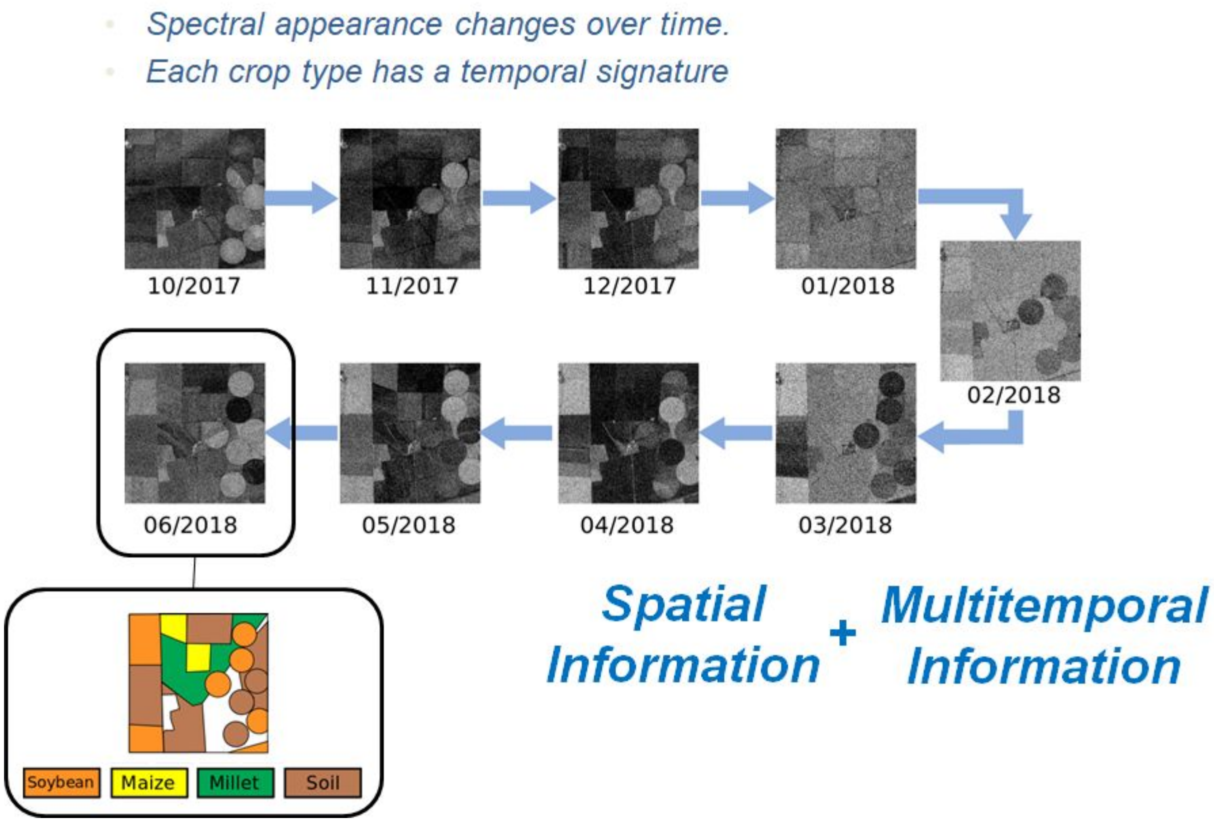


Image
Label
Closed Set
SoftMax^T
OpenFCN
OpenIPCS
UUC: Imp. Surf.
TPR = 0.7
UUC: Building
TPR = 0.7
UUC: Low Veg.
TPR = 0.3
UUC: Tree
TPR = 0.4
UUC: Car
TPR = 0.4
Legend
Impervious Surfaces
Building
Low Vegetation
Tree
Car
Unknown

Open Set Semantic Segmentation For Multitemporal Crop Recognition

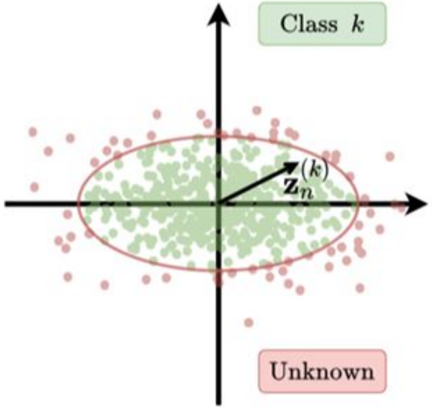


Martinez, J. A. C., Oliveira, H., dos Santos, J. A., & Feitosa, R. Q. (2021). Open Set Semantic Segmentation for Multitemporal Crop Recognition. *IEEE Geoscience and Remote Sensing Letters*.

Open Set Semantic Segmentation For Multitemporal Crop Recognition

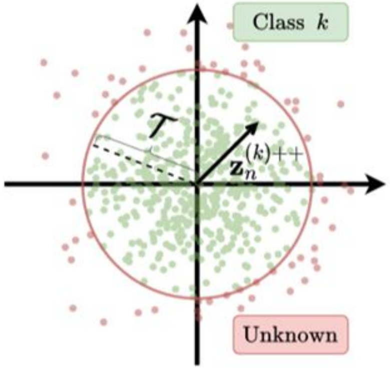
OpenPCS

- Assumed gaussian **<u>distributions are different</u>** for each class: N(0, **S**_k)
- Gaussian representations for each class are **different** but it uses the **<u>same threshold</u> for all classes**



OpenPCS++

- Applies **whitening** transformation that **normalizes** the distribution of each class into **isotropic Gaussian** (Same threshold can be used for all classes)



OpenPCA result

$$\mathbf{z}_n^{(k)++} = S_k^{-1/2}\mathbf{z}_n^{(k)}$$

Normalization (whitening) term

Martinez, J. A. C., Oliveira, H., dos Santos, J. A., & Feitosa, R. Q. (2021). Open Set Semantic Segmentation for Multitemporal Crop Recognition. *IEEE Geoscience and Remote Sensing Letters*.

Final Remarks