NeurIPS 2018

Vision and Security Technology
UNIVERSITY OF COLORADO **COLORADO SPRINGS**

Reducing Network Agnostophobia

Akshay Raj Dhamija
Dr. Manuel Günther
Dr. Terrance E. Boult

NeurIPS 2018

Vision and Security Technology

UNIVERSITY OF COLORADO **COLORADO SPRINGS**

Reducing Network Agnostophobia

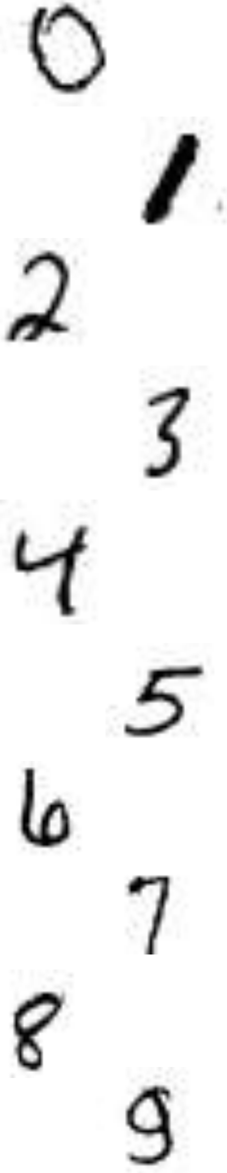
"The Fear of Unknow"

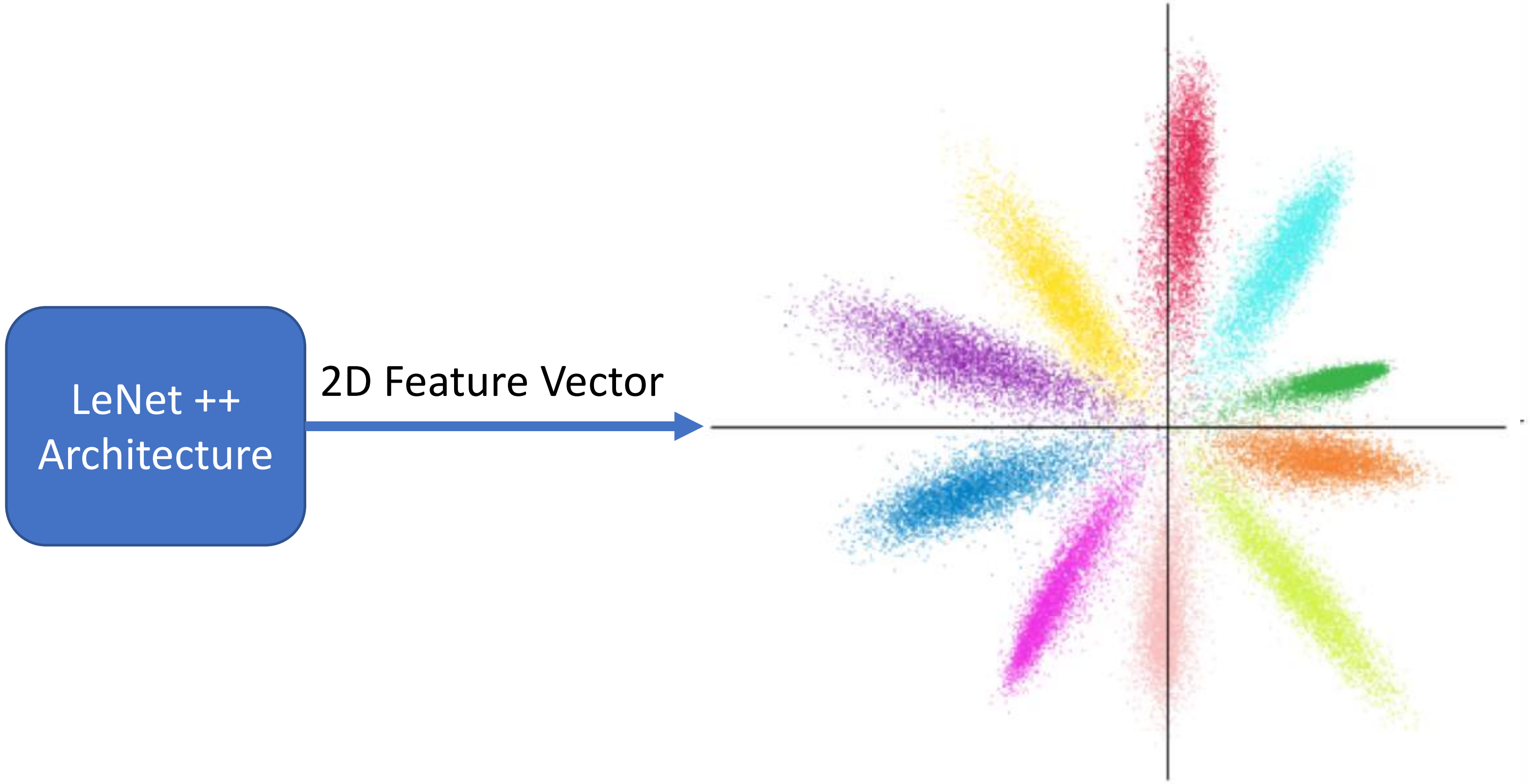
Akshay Raj Dhamija
Dr. Manuel Günther
Dr. Terrance E. Boult

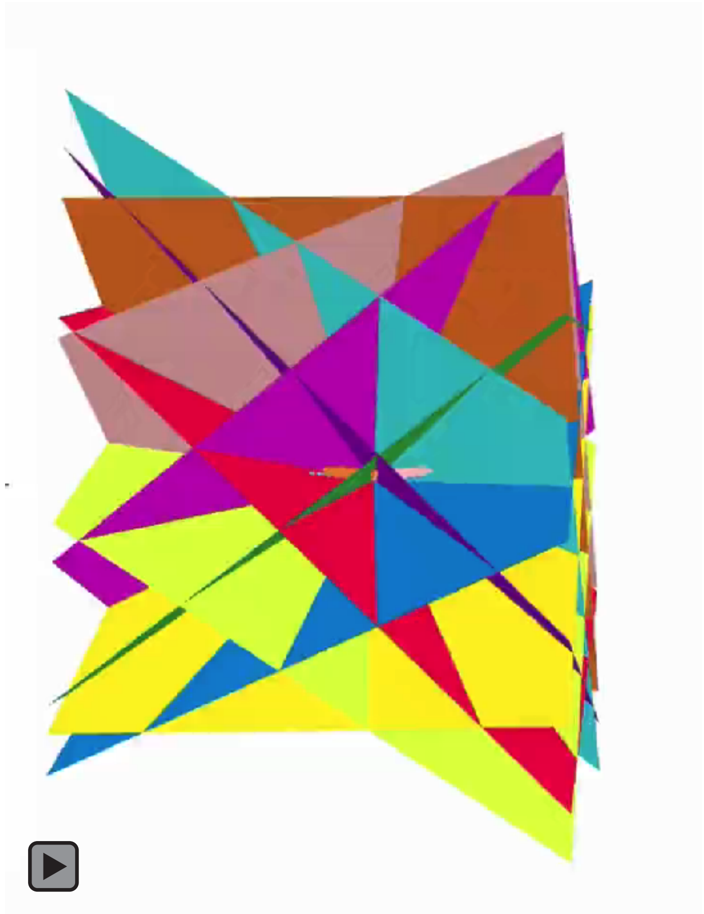
Classification with Deep Neural Networks

Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult
University of Colorado Colorado Springs

Classification with Deep Neural Networks







Deep Feature Representations

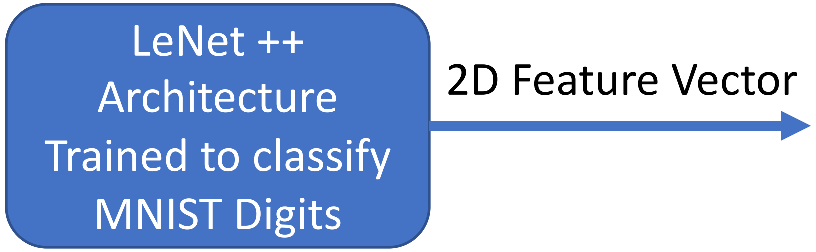
Decision Planes

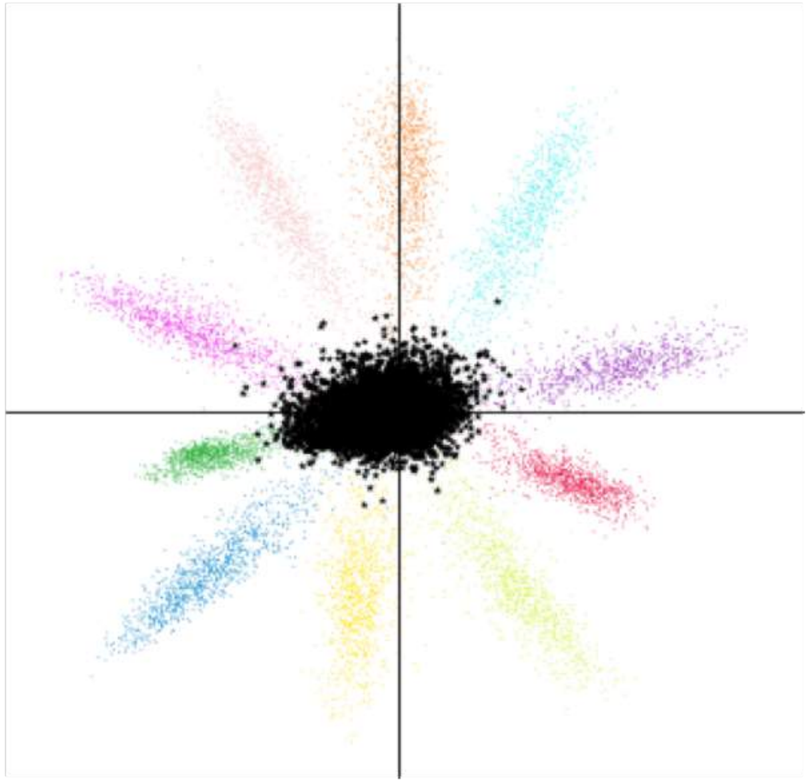
Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult

University of Colorado Colorado Springs

Response to Out of Distribution Samples - CIFAR Samples

Response to Out of Distribution Samples - CIFAR Samples





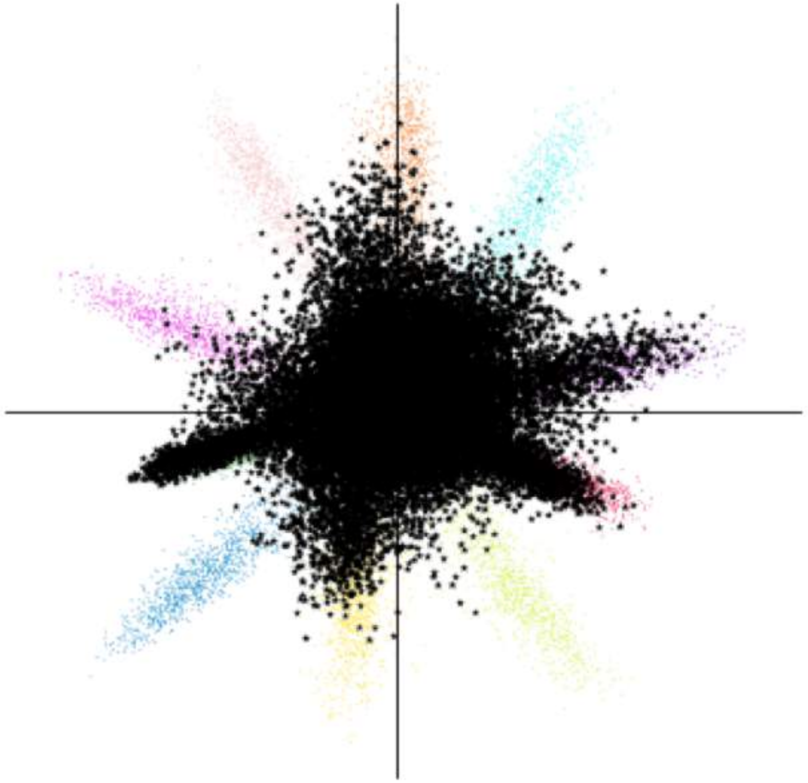
Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult

University of Colorado Colorado Springs

Response to Out of Distribution Samples – NIST Letters

LeNet ++
Architecture
Trained to classify
MNIST Digits

2D Feature Vector



Handwritten Character Recognition Using Neural Network Architectures*

O. Matan, R. K. Kiang, C. E. Stenard, B. Boser,
J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard,
L. D. Jackel, and Y. Le Cun
AT&T Bell Laboratories, Holmdel, N. J. 07733

Abstract

We have developed a neural-network architecture for recognizing handwritten digits. This network has 1% error rate with about 7% reject rate on ... written zipcode digits provided by the U.S. Postal ...vice. In ... architecture ...

One of the earliest approaches for "none of the above" or "none of known classes" – 1990

outputs ... held co...
are other interesting theoretical qualities of Softmax, such as its connection to the entropy of the system (Bridle, 1989). The form of Softmax is the following:

$$S_i = \frac{e^{\beta O_i}}{\sum_k e^{\beta O_k}}$$

Where O_i is the activation level of output unit i, and S_i is the Softmax score for class i. We have slightly modified this function by adding an additional term to the denominator:

$$S_i = \frac{e^{\beta O_i}}{e^{\alpha} + \sum_k e^{\beta O_k}}$$

The term involving α essentially represents the activation level of an artificial N+1st category, the "none of the above" category. It will cause reduction of the score when the highest active unit has a low absolute value.

Softmax considers competition between the most-active unit and all the ...

Handwritten Character Recognition Using Neural Network Architectures*

O. Matan, R. K. Kiang, C. E. Stenard, B. Boser,
J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard,
L. D. Jackel, and Y. Le Cun
AT&T Bell Laboratories, Holmdel, N. J. 07733

Abstract

We have developed a neural-network architecture for recognizing handwritten digits. This network has 1% error rate with about 7% reject rate on ...ritten zipcode digits provided by the U.S. Postal ...vice. In ... architecture ...

One of the earliest approaches for "none of the above" or "none of known classes" – 1990

outputs ... held co... are other interesting theoretical qualities of Softmax, such as its connection to the entropy of the system (Bridle, 1989). The form of Softmax is the following:

$$S_i = \frac{e^{\beta O_i}}{\sum_k e^{\beta O_k}}$$

Standard Softmax

Where O_i is the activation level of output unit i, and S_i is the Softmax score for class i. We have slightly modified this function by adding an additional term to the denominator:

$$S_i = \frac{e^{\beta O_i}}{e^{\alpha} + \sum_k e^{\beta O_k}}$$

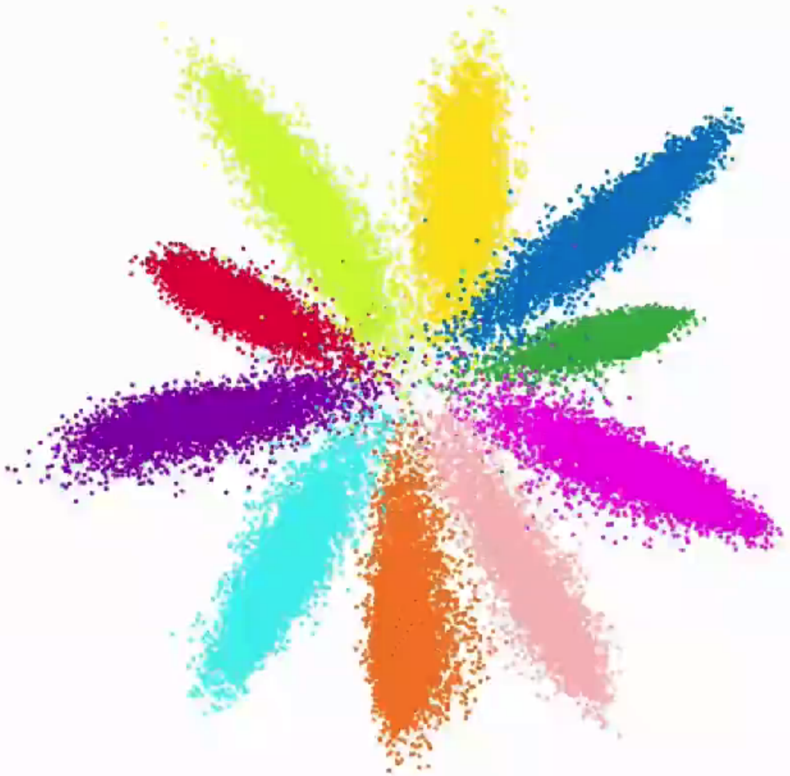
Modified Softmax

The term involving α essentially represents the activation level of an artificial N+1st category, the "none of the above" category. It will cause reduction of the score when the highest active unit has a low absolute value.

Softmax considers competition between the most-active unit and all the ...

Side View

Top View



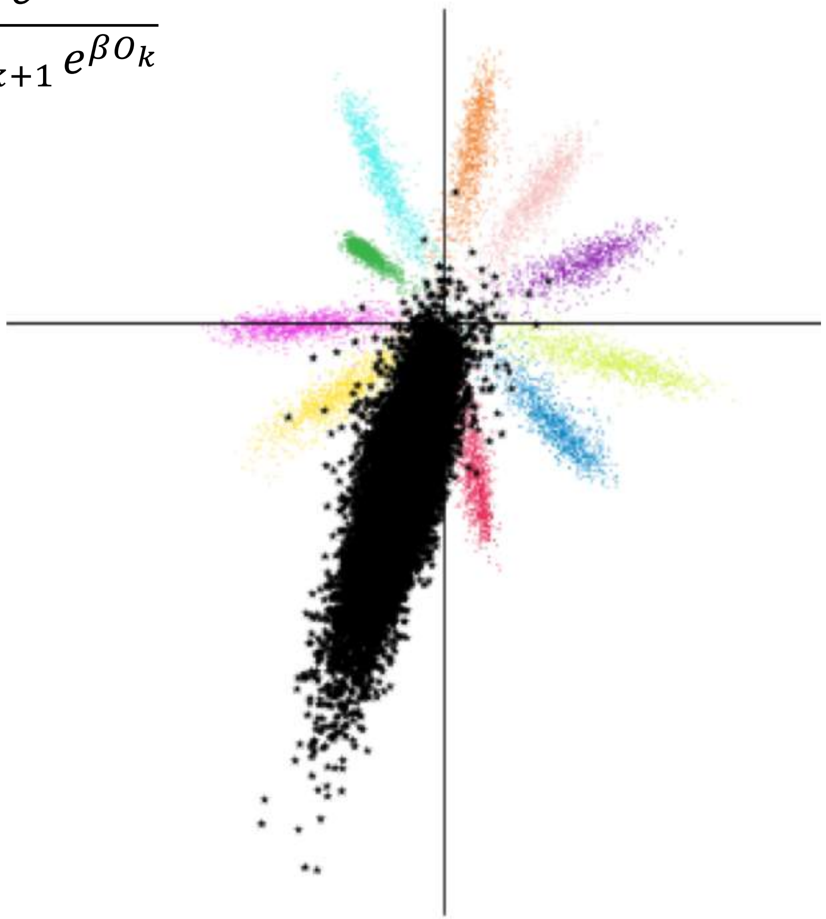
Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult

University of Colorado Colorado Springs

Background Class Approach

$$S_i = \frac{e^{\beta o_i}}{e^{\alpha} + \sum_k e^{\beta o_k}} \longrightarrow S_i = \frac{e^{\beta o_i}}{\sum_{k+1} e^{\beta o_k}}$$

Learning α plane
and better features to
separate the unknowns



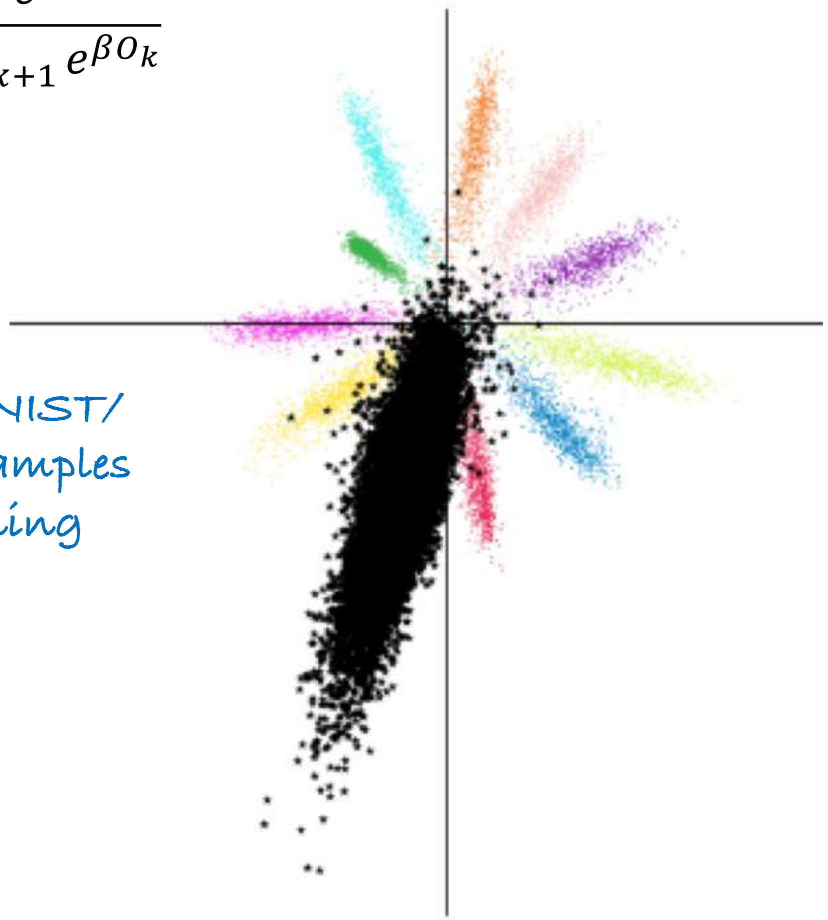
Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult
University of Colorado Colorado Springs

Background Class Approach

$$S_i = \frac{e^{\beta o_i}}{e^{\alpha} + \sum_k e^{\beta o_k}} \longrightarrow S_i = \frac{e^{\beta o_i}}{\sum_{k+1} e^{\beta o_k}}$$

Learning α plane
and better features to
separate the unknowns

Pool of non-MNIST/
Background samples
during training



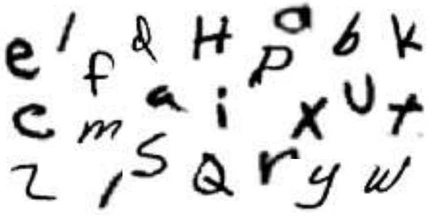
Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult

University of Colorado Colorado Springs

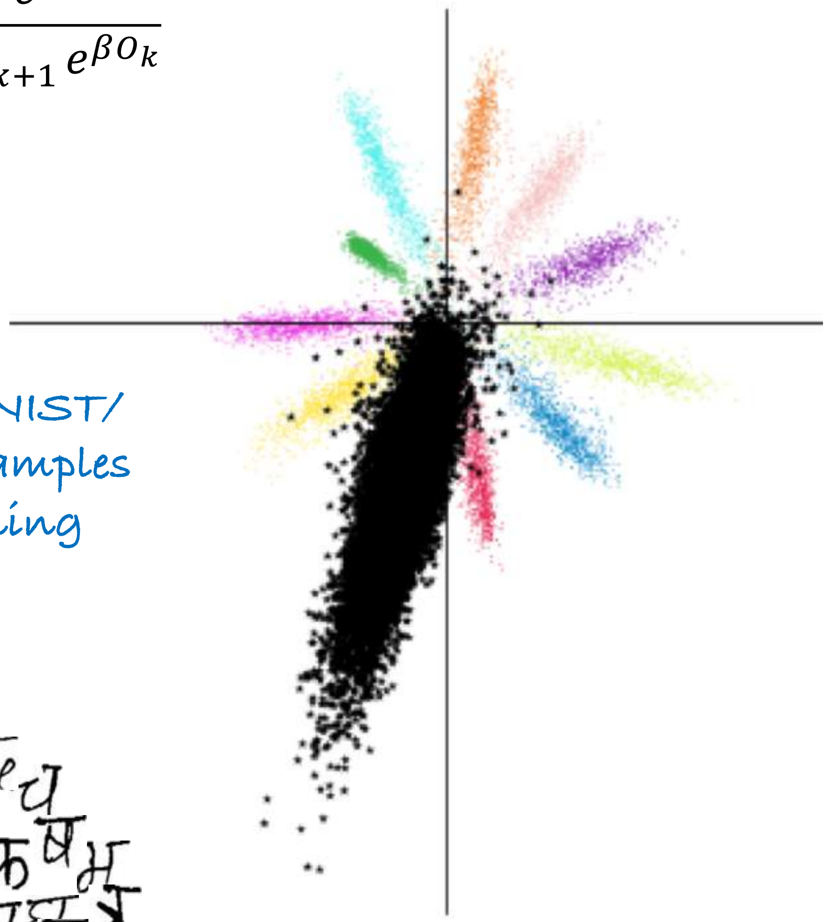
Background Class Approach

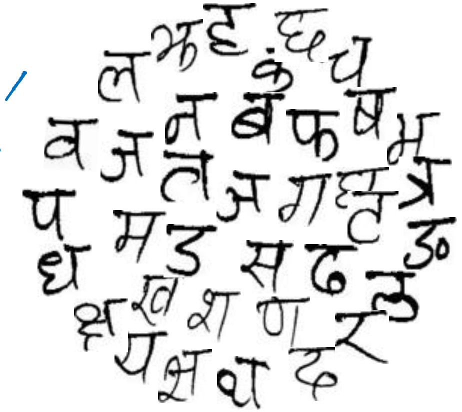
$$S_i = \frac{e^{\beta o_i}}{e^{\alpha} + \sum_k e^{\beta o_k}} \rightarrow S_i = \frac{e^{\beta o_i}}{\sum_{k+1} e^{\beta o_k}}$$

Learning α plane
and better features to
separate the unknowns



Pool of non-MNIST/
Background samples
during training



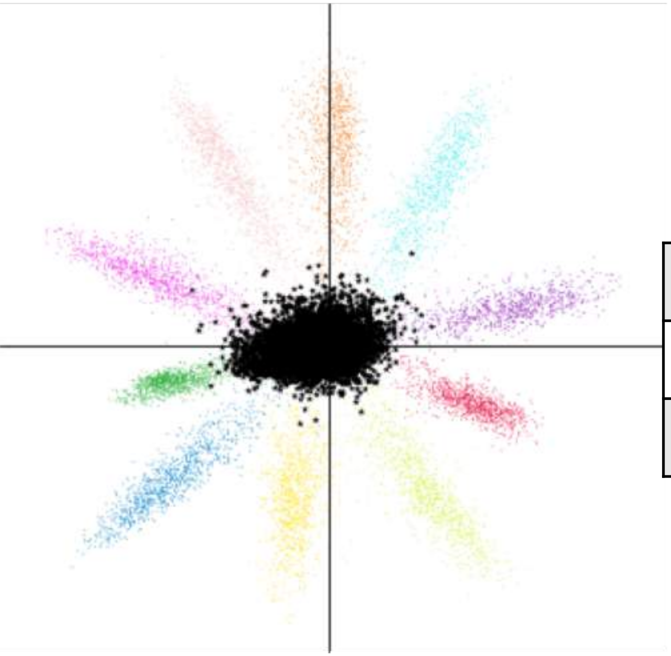


Out-of-distribution /
Unkonwn samples
for testing

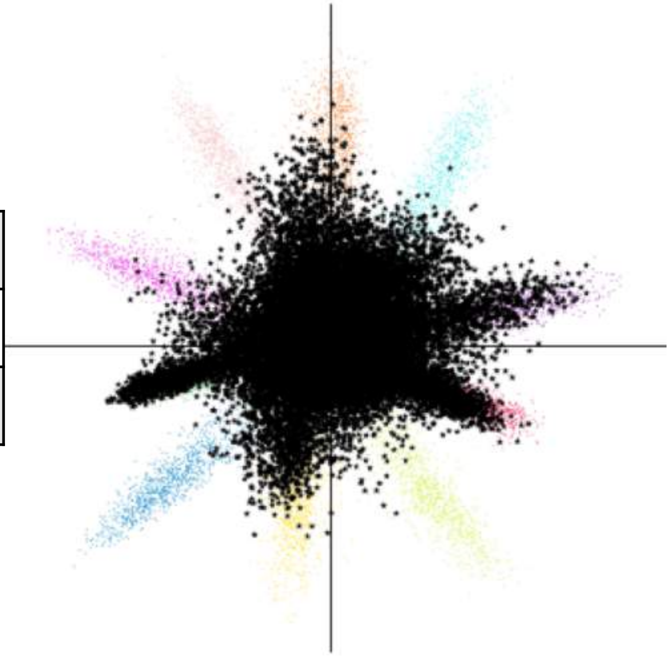
Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult

University of Colorado Colorado Springs

Observation from Default Response – Leading to Our Approach

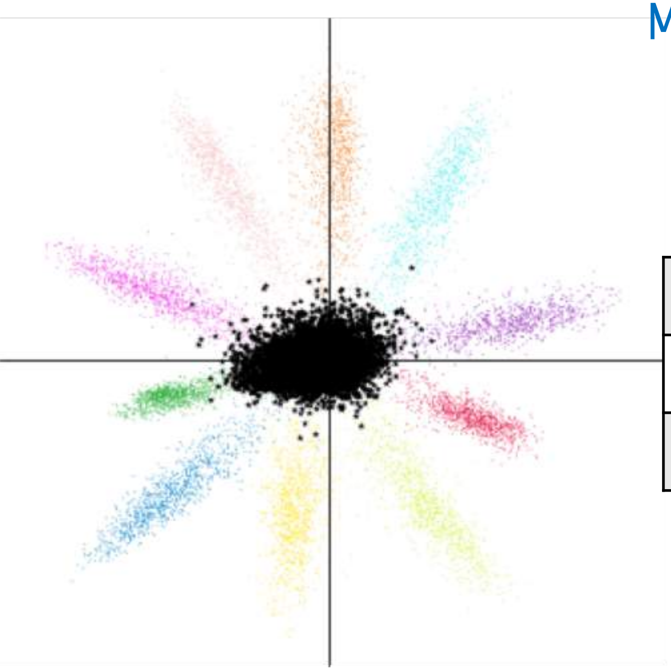


Magnitude of Feature Vector		Entropy	
Knowns	Unknowns	Knowns	Unknowns
94.90 ± 27.47	32.27 ± 18.47	0.015 ± .084	0.318 ± .312



Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult
University of Colorado Colorado Springs

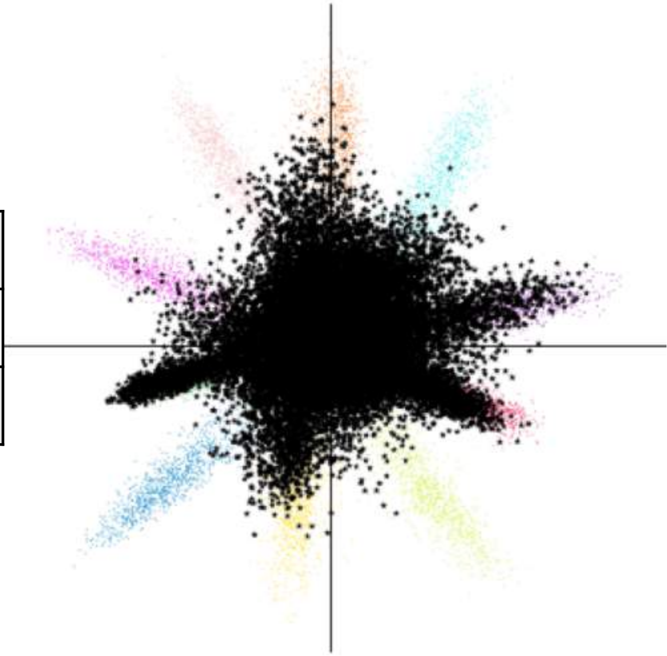
Observation from Default Response - Leading to Our Approach



Magnitude of Deep Feature Representations of

Known Samples > Unknown Samples

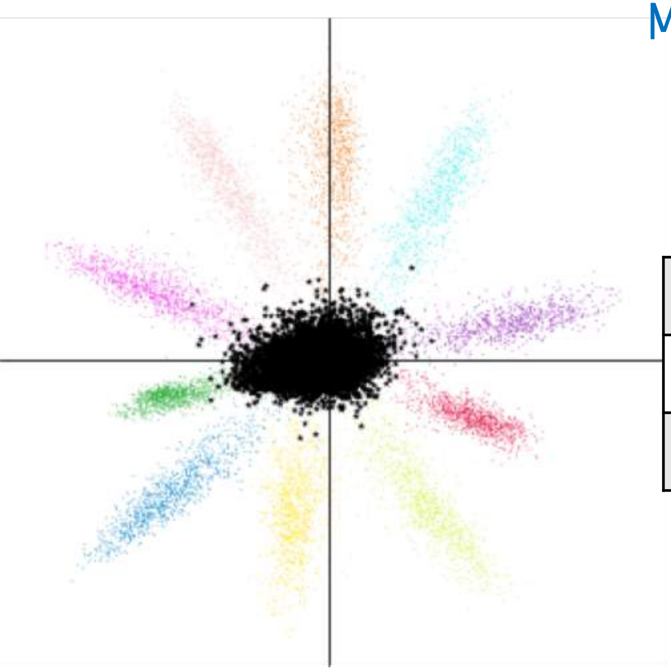
Magnitude of Feature Vector		Entropy	
Knowns	Unknowns	Knowns	Unknowns
94.90 ± 27.47	32.27 ± 18.47	0.015 ± .084	0.318 ± .312



Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult

University of Colorado Colorado Springs

Observation from Default Response - Leading to Our Approach



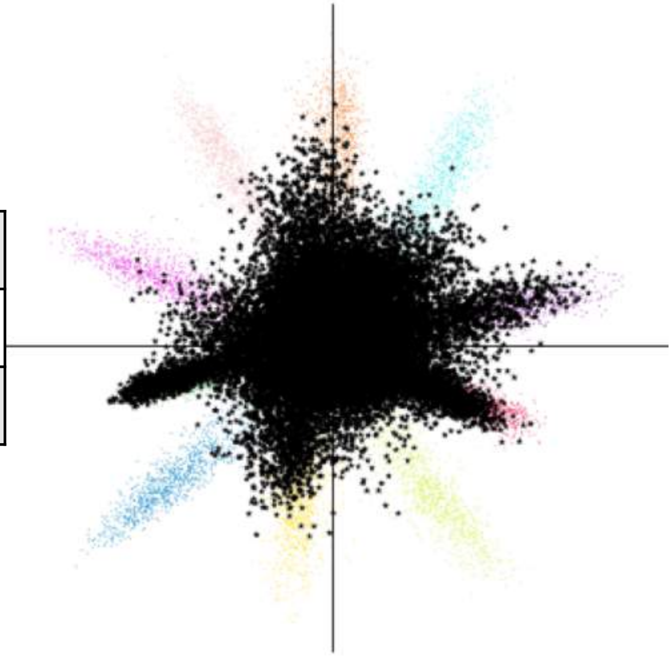
Magnitude of Deep Feature Representations of

Known Samples > Unknown Samples

Magnitude of Feature Vector		Entropy	
Knowns	Unknowns	Knowns	Unknowns
94.90 ± 27.47	32.27 ± 18.47	0.015 ± .084	0.318 ± .312

Entropy of

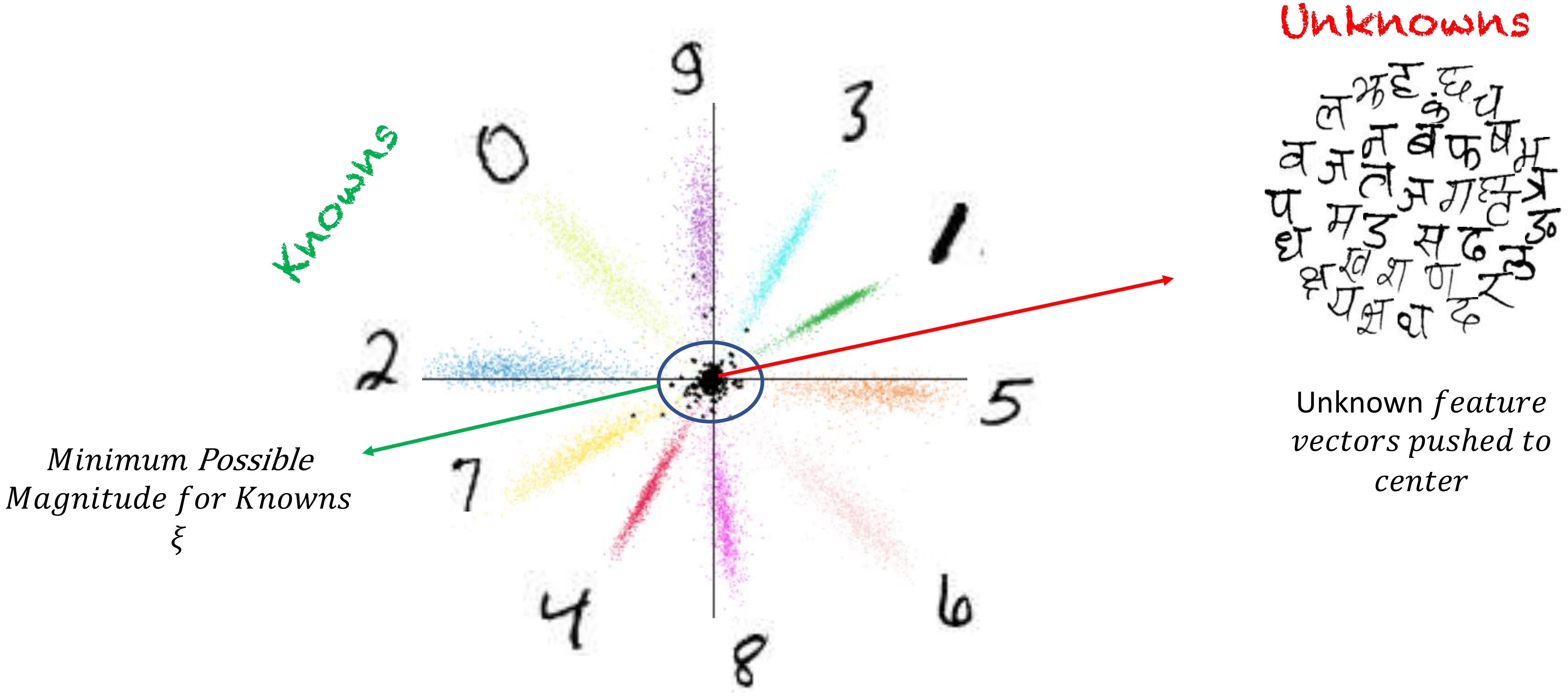
Known Samples < Unknown Samples



Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult

University of Colorado Colorado Springs

Our Approach



Our Approach

Entropic Open-Set Loss

$$\mathcal{J}_E(x) = \begin{cases} -\log S_c(x) & \quad if\ x \in\ D_c' \\ -\dfrac{1}{C}\displaystyle\sum_{c=1}^{C} \log S_c(x) & \quad if\ x \in\ D_b' \end{cases}$$

Increase Entropy
Margin

Increases entropy of the softmax scores for unknwons

Our Approach

Entropic Open-Set Loss

$$\mathcal{J}_E(x) = \begin{cases} -\log S_c(x) & if\ x \in\ D'_c \\ -\dfrac{1}{C}\displaystyle\sum_{c=1}^{C} \log S_c(x) & if\ x \in\ D'_b \end{cases}$$

Increase Entropy Margin

Increases entropy of the softmax scores for unknwons

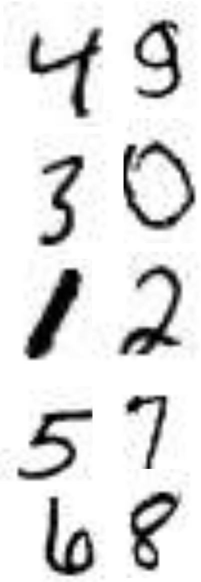
Objectosphere Loss

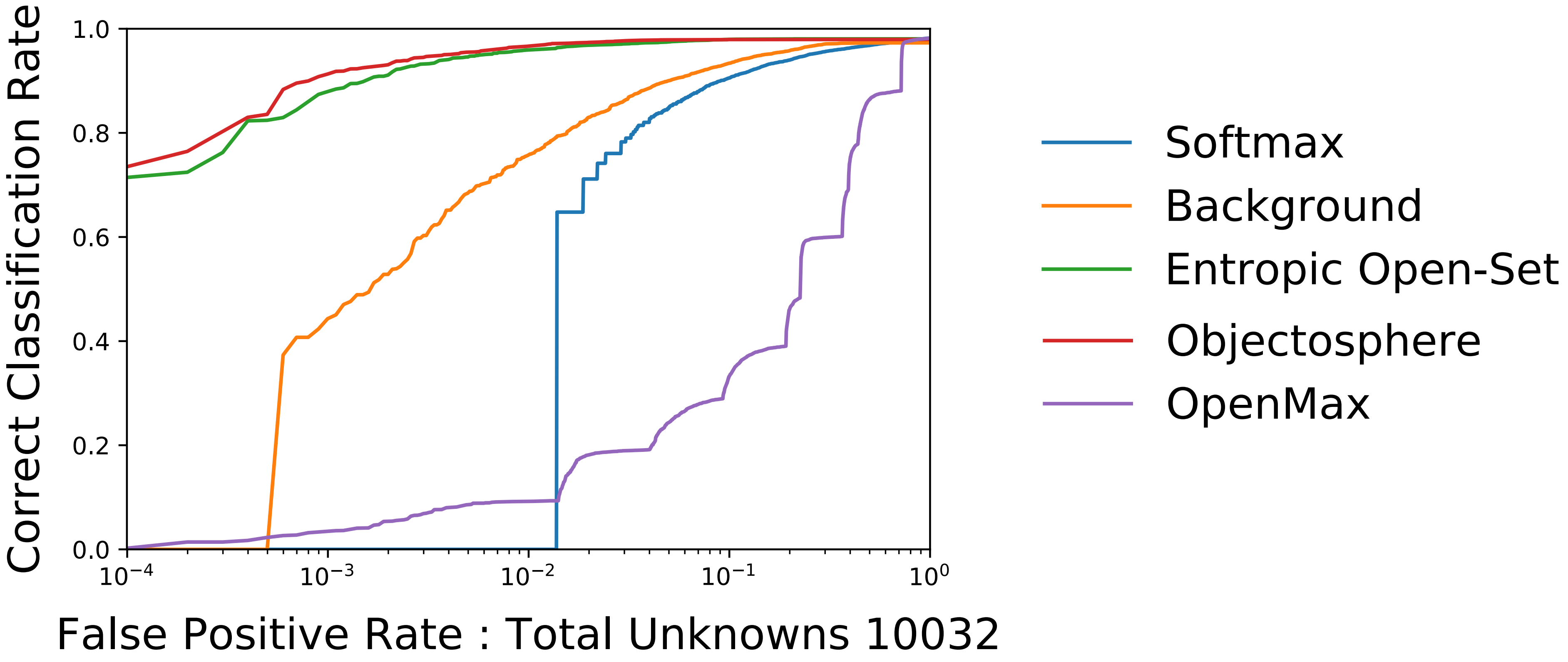
Increase Deep Feature Magnitude Margin

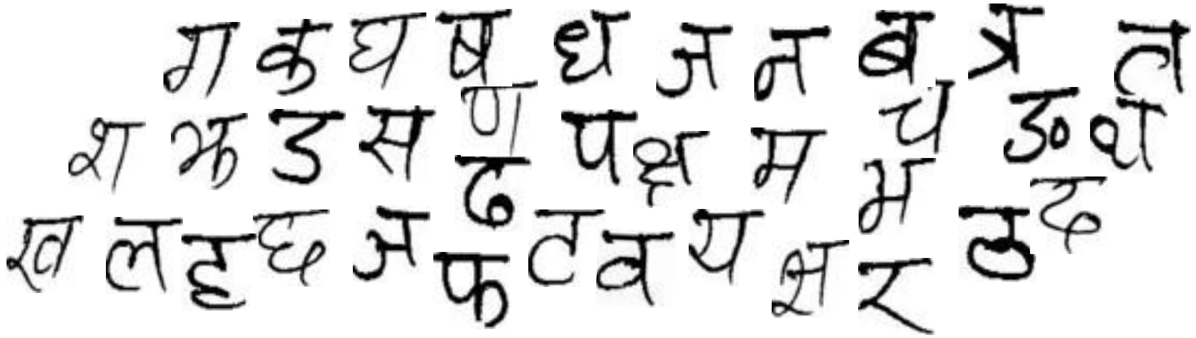
$$\mathcal{J}_R(x) = \ \mathcal{J}_E + \begin{cases} \max(\xi\ - \| \ F(x)\ \|, 0)^2 - \log S_c(x) & if\ x \in\ D'_c \\ \| \ F(x)\ \|^2 & if\ x \in\ D'_b \end{cases}$$

Minimizing the Euclidean length of deep representations for unknonwns

Novel Evaluation Metric : Open Set Recognition Curve







Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult

University of Colorado Colorado Springs

Thank You!

What's at the Poster B#100 ?

- Why this works!
- Drawbacks of current evaluation techniques
- Discussion of Openset deep networks
- Performance on wider networks like ResNet-18
- Implementing Entropic Open-Set loss in 1-line of code ☺

Thoughts and comments are welcome at adhamija@vast.uccs.edu

 Akshay Raj Dhamija | Dr. Manuel Günther | Dr. Terrance E. Boult

University of Colorado Colorado Springs