Dual-Tree Wavelet Scattering Network with Parametric Log Transformation for Object Classification

Amarjot Singh and Nick Kingsbury- University of Cambridge, U.K.

UNIVERSITY OF CAMBRIDGE

Problem Objective: Object Classification

- *Deep Models* (CNNs) give state-of-the-art object classification accuracy but are not well understood and difficult to optimize for small training datasets.
- *ScatterNet* [1] is a mathematical framework that extracts invariant features which give similar classification performance to learned networks on some datasets.
- We propose a *computationally efficient DTCWT wavelet ScatterNet with Parametric Log Transformation* that extracts relatively symmetric translation invariant features.

Contributions

The contributions of the proposed ScatterNet are:

- *Multi-resolution Input Images:* The input image is transformed into multi-resolution images of different sizes such that the dual-tree wavelet decompositions produce densely spaced feature maps over scale.
- *Parametric Log Transformation:* Log transformation reduces the effect of outliers by introducing approximate symmetry in extracted features. The transformation also creates a form of contrast normalization which enhances weaker features.
- *Computational Efficiency:* Dual-Tree wavelets are used as opposed to Morlet because they extract features with less computations. They also have the properties of perfect reconstruction and limited redundancy.

DTCWT ScatterNet Formulation

The formulation that captures translation invariant features from a single image is presented. This can be extended to each of the multi-resolution (R1, R2) images.

A. *Translation Invariant relatively symmetric features*:

$$\widetilde{W}_1 x = (x \star \phi_J, |x \star \psi_{\theta,j}|_{\theta,j}) = (S_0 x, Ux)$$

$$U[\lambda_{m=1}] = |x \star \psi_{\lambda_1}| = \sqrt{|x \star \psi^a_{\lambda_1}|^2 + |x \star \psi^b_{\lambda_1}|^2}$$

$$U1[j] = \log(U[j] + k_j), \quad U[j] = |x \star \psi_j|$$

$$S[\lambda_{m=1}] = |U1[\lambda_{m=1}]| \star \phi_{2^J}$$

- *Information lost due to smoothing is recovered:*

$$U2[\lambda_{m=1}, \lambda_{m=2}] = |U1[\lambda_{m=1}] \star \psi_{\lambda_{m=2}}|$$

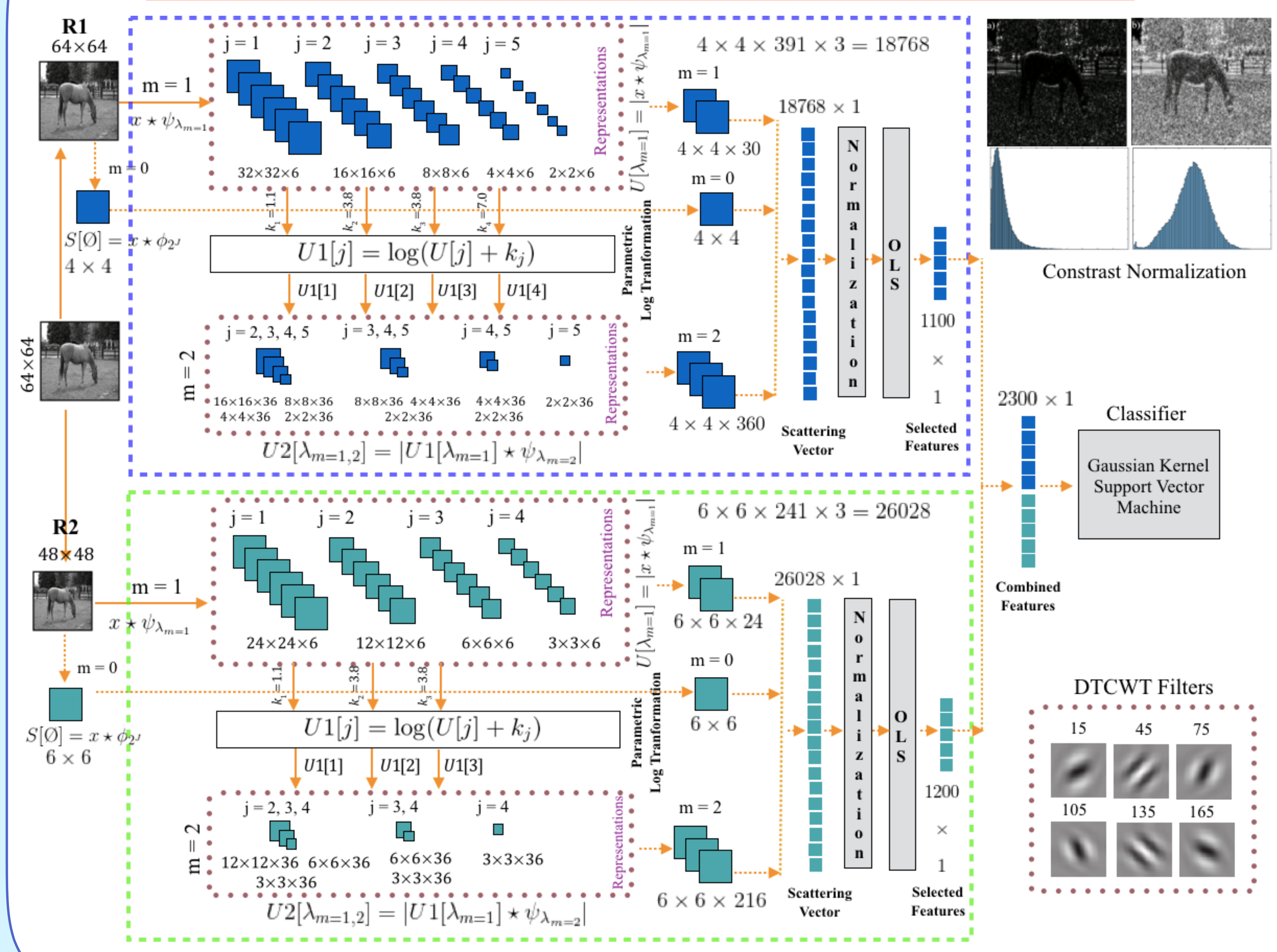
$$S[\lambda_{m=1}, \lambda_{m=2}] = U2[\lambda_{m=1}, \lambda_{m=2}] \star \phi_{2^J}$$

$$S_{R1} = (S_0 x, S[\lambda_{m=1}], S[\lambda_{m=1}, \lambda_{m=2}])$$

B. *Feature selection and classification with G-SVM*:

$$OLS(S_{R1} + S_{R2}) \rightarrow \boxed{\text{Gaussian-Kernel SVM}}$$

DTCWT ScatterNet Architecture



Experimental Results

- Comparison with Mallat's ScatterNet [1], Unsupervised (Unsup) and supervised (Sup) learning methods

Dataset	Pro.	ScatNet [1]	Unsup	Sup
CIFAR-10	**82.4**	81.6	82.2 [12]	89.6 [6]
CIFAR-100	**56.7**	55.8	54.2 [4]	64.3 [6]

- Computationally efficient as compared to Mallat's ScatterNet [1].

Arch.	FVL	SD	FR (%)	TS (s)	T-OLS (h)
ScatNet [1]	113712	2000	1.75	0.98	3.22
R1	18762	1100	**5.86**	**0.46**	**1.07**
R2	26028	1200	**4.61**	**0.32**	**1.14**
Pro. (R1+R2)	44796	2300	**5.13**	**0.78**	**2.21**

- Outperforms LeNet (LN) and Network in Network (NIN) supervised learning methods on small datasets.

Arch.	300	500	1K	2K	5K	10K	20K	50K
Pro.	**39.3**	**48.8**	**55.9**	**61.8**	**67.0**	**72.9**	76.8	82.4
LN	34.9	44.7	53.1	57.9	63.0	69.0	74.0	77.6
NIN	10.1	10.3	10.9	40.4	63.4	72.0	**83.1**	**89.6**

Conclusions

- Proposed DTCWT ScatterNet *outperforms* Mallat's ScatterNet [1] on classification accuracy and computational efficiency.
- DTCWT ScatterNet gives *superior* classification accuracy over the unsupervised learning method.
- The proposed ScatterNet *outperforms* supervised learning methods for small training datasets.

References

1. E. Oyallon, Mallat, S.: Deep Roto-Translation Scattering for Object Classification. CVPR, 2015.