Quaternion Capsule Networks

Barış Özcan[1] Furkan Kınlı[2] Furkan Kıraç[3]

Department of Computer Science, Özyeğin University, İstanbul, Türkiye

[1]baris.ozcan.10097@ozu.edu.tr, {[2]furkan.kinli, [3]furkan.kirac}@ozyegin.edu.tr

Abstract

Capsules are grouping of neurons that allow to represent sophisticated information of a visual entity such as pose and features. In the view of this property, Capsule Networks outperform CNNs in challenging tasks like object recognition in unseen viewpoints. In this paper, we present Quaternion Capsules (QCN) where pose information of capsules and their transformations are represented by quaternions. Quaternions are immune to the gimbal lock, have straightforward regularization of the rotation representation for capsules, and require a smaller number of parameters than matrices. The experimental results show that QCNs generalize better to novel viewpoints with fewer parameters, and also achieve on-par or better performances with the state-of-the-art Capsule architectures on well-known benchmarking datasets.

Quaternion Capsules

$$\hat{v}_{j|i} = f_ij(u_i) = w_{ij} * u_i * w_{ij}^*$$

$$w_{ij} = \left[\cos\theta_{ij}, \quad \sin\theta_{ij}\frac{\bar{\mathbf{w}}_{\mathbf{ij}}}{\|\bar{\mathbf{w}}_{\mathbf{ij}}\|}\right]$$

$$\bar{\mathbf{w}} \;=\; [w_1 i \quad w_2 j \quad w_3 k]$$

- u_i is the child capsule pose and w_{ij} is the learned transformation between i^{th} and j^{th} capsule in layers L and $L+1$ respectively.
- w_{ij}^* stands for the conjugate and $*$ stands for quaternion product.
- $\bar{w}$ is the rotation axis and θ is the rotation angle (Figure 1).
- EM Routing is used without any modification for fair comparison.

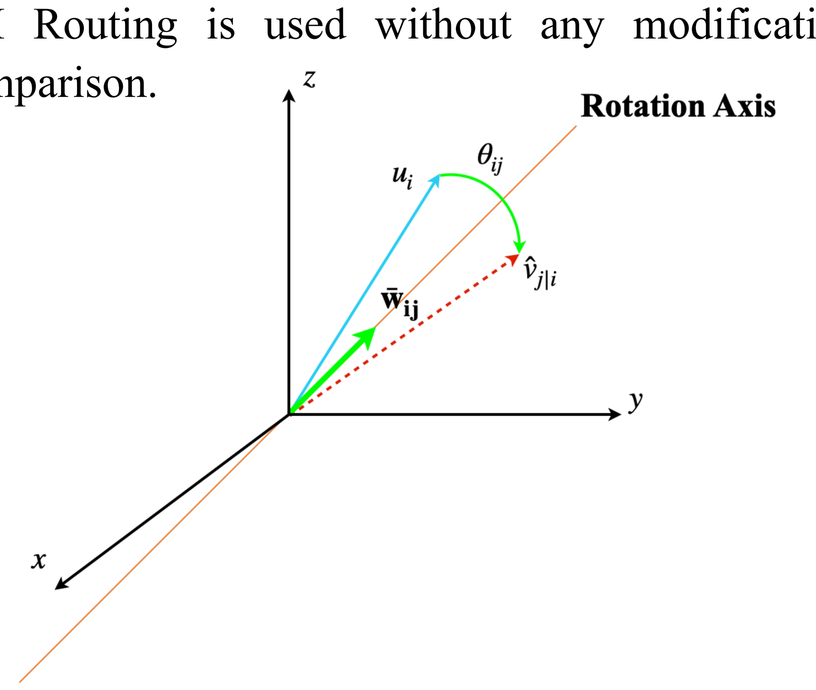

Figure 1. Rotation of a capsule pose to estimate its parents pose via quaternions.

Our Architecture

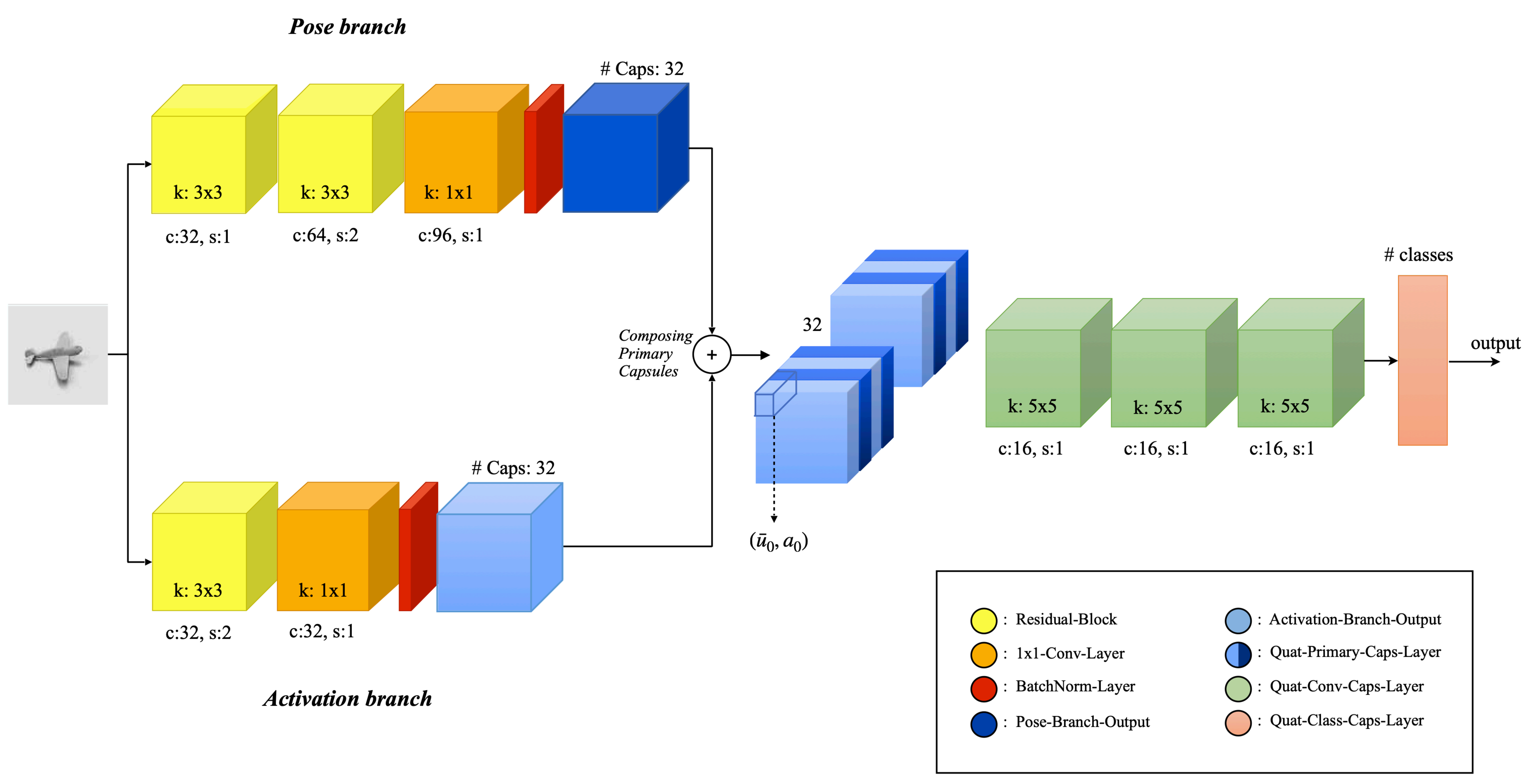

Figure 2. Our two branched architecture for Quaternion Capsule Networks.

Effect of Branching

- Experiments on smallNORB with branched and non-branched versions of QCN and Matrix Capsules.
- Branched version of Matrix Capsules is consist of the same capsule extractor layers as in QCN, while non-branched versions have two consecutive residual blocks with 64 and 96 channels.
- The results demonstrate that using separate branches for pose and activation increases the overall performance of QCN, though the number of parameters is reduced.

Table 1. The effect of branching on the error rates and parameters for both Matrix and Quaternion Capsules.

Models	Error Rate (%)			~# Parameters		
	with	without	diff.	with	without	diff.
EM‡	3.66	3.44	0.22	411K	533K	-122K
QCN	2.29	3.72	**-1.43**	188K	298K	-110K

Results on Common Datasets

Table 2. Results on Common datasets where capsule networks are tested.

Models	smallNORB		MNIST		FashionMNIST		SHVN		CIFAR-10	
	Error (%)	# of Params	Error (%)	# of Params	Error (%)	# of Params	Error (%)	# of Params	Error (%)	# of Params
EM [7]	1.8	~310K	0.44	–	–	–	–	–	11.9	–
EM-IBM [37]	4.6	~335K	1.23‡	~337K‡	10.44‡	~337K‡	–	–	–	–
VB [8]	1.6	~169K	–	–	5.2	~172K	3.9	~323K	11.2	~323K
EM*	3.40	~317K	0.89	~319K	9.74	~319K	8.19	~320K	17.76	~460K
QCN	2.29	~188K	0.37	~187K	6.92	~187K	4.63	~189K	13.92	~189K

Generalization to Novel Viewpoints

- Training and test samples are from different viewpoints of the classes in smallNORB dataset.
- Azimuth experiment: training and test set are divided w.r.t. azimuth angles
- Elevation experiment: training and test set are divided w.r.t. elevations
- Training is stopped where training performances are similar for each model to test generalization fairly.

Train samples Test samples

Table 3. Generalization to novel viewpoint results. Methods are trained on Familiar viewpoints and tested on novel viewpoints.

Viewpoints	Azimuth (%)				
(Models)	QCN	EM*	VB	EM	CNN
Novel	**7.5**	13.4	11.3	13.5	20.0
Familiar	3.7	3.7	3.7	3.7	3.7
Viewpoints	**Elevation (%)**				
(Models)	QCN	EM*	VB	EM	CNN
Novel	**11.5**	15.8	11.6	12.3	17.8
Familiar	4.4	4.0	4.3	4.3	4.3

- EM*: our implementation of [1].
- EM : reported results of [1].
- VB: Capsule Routing via Variational Bayes [2].
- CNN: Baseline reported result in [1].

References

[1] Hinton, Geoffrey E., Sara Sabour, and Nicholas Frosst. "Matrix capsules with EM routing." International conference on learning representations. 2018.

[2] Ribeiro, Fabio De Sousa, Georgios Leontidis, and Stefanos D. Kollias. "Capsule Routing via Variational Bayes." *AAAI*. 2020.

[3] Trabelsi, C., et al. "Deep Complex Networks. arXiv 2018." *arXiv preprint arXiv:1705.09792*.[