





## **Connectivity Learning in Multi-Branch Networks** Karim Ahmed, Lorenzo Torresani

## Extended paper: arXiv:1709.09582



	r
	Results
m data by	<b>CIFAR-100</b> (*
Learned binary masks	Effect o
ine connectivity for each block	72
$\begin{array}{c} g_{j}^{(i-1)} \\ 64, 1X1, 4 \\ 64, 1X1, 4 \\ 4, 3X3, 4 \\ 4, 1X1, 64 \end{array}$	<b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b> <b>AC</b>
	6/ K=1 K=2 K=3 K=4 Number of active co
	ArchitectureCo $\{Depth (D), Bottleneck width (w), Cardinality (C)\}$
Our Approach	Fix {29,8,8} Fix Le
	Le Fiz
gation defines active	{29,64,8} Fix Le Le
connections	<b>CIFAR-10</b> (10
	Architecture
	$\left\{ \text{Depth}(D), \text{Bottleneck width}(w), \text{Cardinality}(C) \right\}$
	$\{20,\!4,\!8\}$
ut connections (fan-in)	{29,8,8}
anyporparanetty	{29,64,8}
C into $\mathbf{a}^{(i)} \subset [0, 1]^C$ of $\sum_{i=1}^{C} a^{(i)} = K$	ImageNet (10
$ \text{INTO } \mathbf{g}_j \subset \{0, 1\}  \text{S.t. } \sum_{k=1}^{g_{j,k}} g_{j,k} = \mathbf{n} $	Architecture $\langle Depth(D) \rangle$ Bottleneck w
y masks $\mathbf{g}_j^{(i)} \in \{0,1\}^C$	Cardinality $(C)$
	{50,4,32}
	{101,4,32}
	{101,4,64}
	References
	Xie et al. "Aggregated residu
ued masks $\tilde{g}_{ik}^{(i)}$	Szegedy et al. "Rethinking the
$\mathcal{I}_{J}, \mathcal{N}$	Courbariaux, et al. "Binarycon
	during propagations." NIPS 2