ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

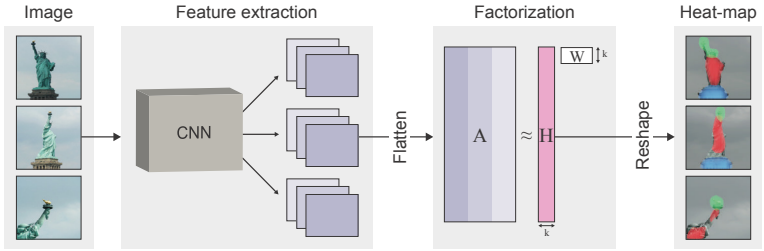
Deep Feature Factorization for Concept Discovery

Edo Collins, Radhakrishna Achanta, Sabine Süsstrunk

ivrl.epfl.ch

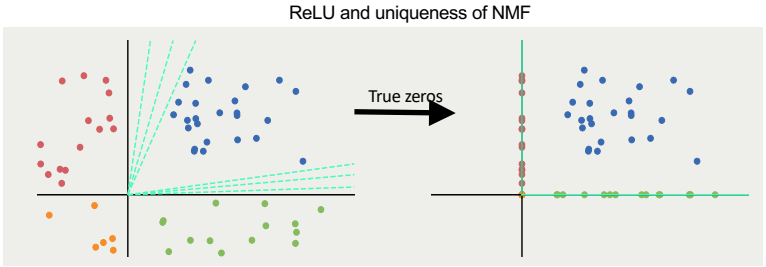
Deep Feature Factorization for Concept Discovery

Which objects are shared across images? **DFF allows to see how a deep CNN perceives similarities** in an image set by producing heat maps that localize semantic objects and object-parts, shown on the right.



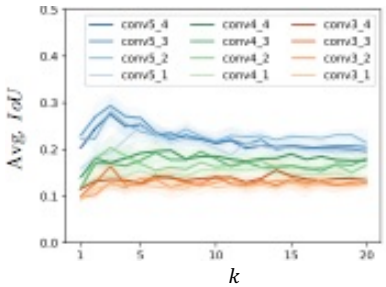
Deep Feature Factorization (DFF) is the application of **non-negative matrix factorization** to the feature activation of a deep neural network. In the case of deep convolutional networks, such as VGG-19 used here, each of the k DFF factors defines a heat map over the image batch.

DFF exploits the **geometry of rectified-linear units** (ReLU), ubiquitous across most deep neural architectures. ReLU setting negative values to true zeros (see below) creates favorable conditions for NMF in deep feature space, since the estimation problem $A \approx H_+ W_+$ becomes less ill-posed as $\|A\|_0$ decreases, since $A_{i,j} = 0 \rightarrow \exists_k \, s.t. \, H_{i,k} = 0$.

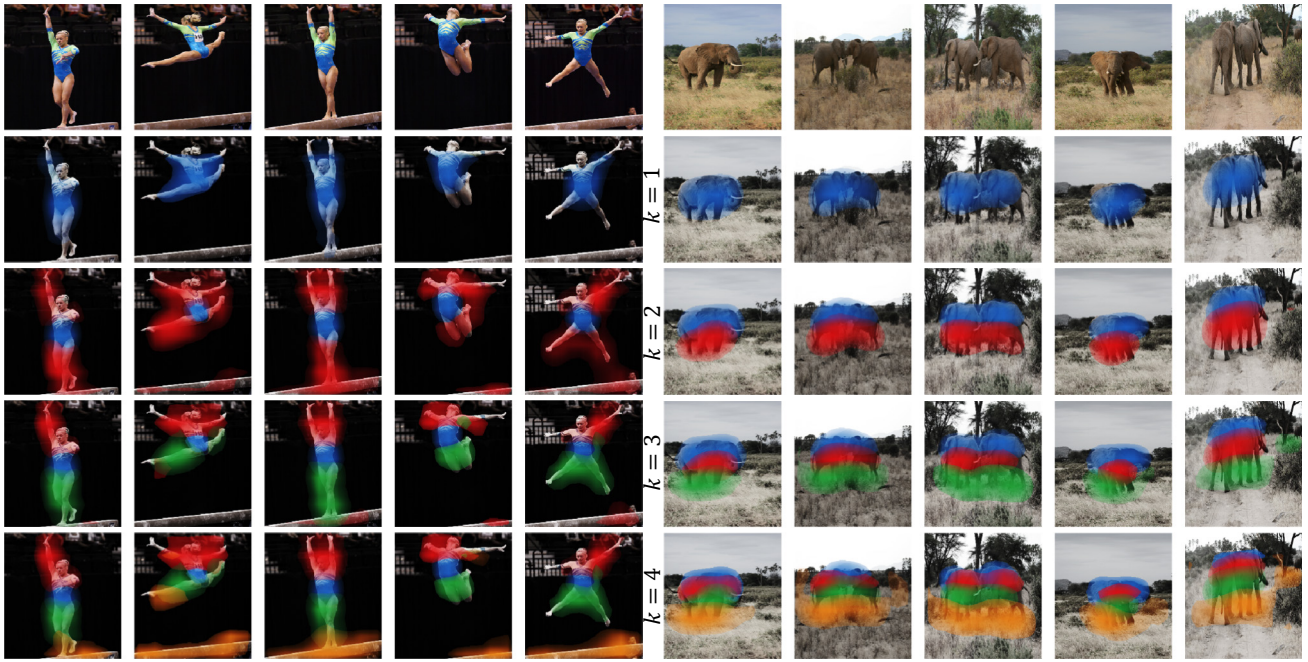


The correspondence to semantic parts is depth dependent, as shown over the layers of VGG-19.

The results shown are for a subset of the iCoseg dataset, which we further annotated with pixel-level part labels.



Applying DFF with increasing k reveals **a concept hierarchy**, e.g., the cluster qualitatively corresponding to *body* is split into *limbs* and *torso*, and then *limbs* is further split into *arms* and *legs.* DFF shows **invariance to complex transformations,** such as the varied leg positions of the gymnast on the left and back side of the elephants on the right.



DFF with $k = 1$ achieves **state-of-the-art results** on weakly-supervised object co-localization on PASCAL VOC 2007, while being a much simpler method. Numbers indicate CorLoc scores.

Method	aero	bicy	bird	boa	bot	bus	car	cat	cha	cow	dtab	dog	hors	mbik	pers	plnt	she	sofa	trai	tv	mean
Joulin et al.	33	17	21	18	5	27	33	41	6	29	**35**	32	26	40	18	12	25	28	36	12	25.60
Cho et al.	50	43	30	19	4	62	**65**	43	**9**	49	12	44	64	57	15	9	31	34	62	**32**	36.60
Li et al.	**73**	45	43	28	7	53	58	45	6	48	14	47	**69**	67	**24**	13	**52**	26	65	17	40.00
Le et al. (A)	70	52	44	**30**	5	56	60	59	6	49	16	51	59	67	23	12	47	27	59	16	40.36
Le et al. (V)	72	**62**	48	28	**12**	**64**	59	72	6	37	12	45	67	72	19	11	37	29	**67**	23	41.97
DFF	61	49	**54**	20	10	60	46	**79**	4	51	32	**67**	66	70	19	**15**	40	32	66	20	42.87
DFF-CRF	64	47	50	16	10	62	52	75	8	**53**	**35**	65	65	**72**	16	14	41	**36**	63	30	**43.51**