

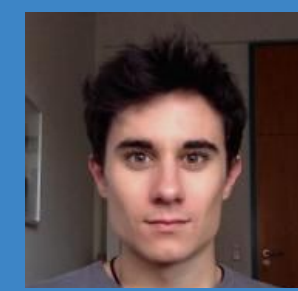
# More Cat than Cute? Interpretable Prediction of Adjective Noun-Pairs



Dèlia Fernández



Alejandro Woodward



Víctor Campos



Brendan Jou



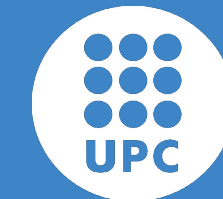
Xavier Giró



Shih-Fu Chang



COLUMBIA UNIVERSITY  
IN THE CITY OF NEW YORK



UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH



Barcelona  
Supercomputing  
Center  
Centro Nacional de Supercomputación



MUSA2 Workshop  
ACM Multimedia 2017

## Motivation

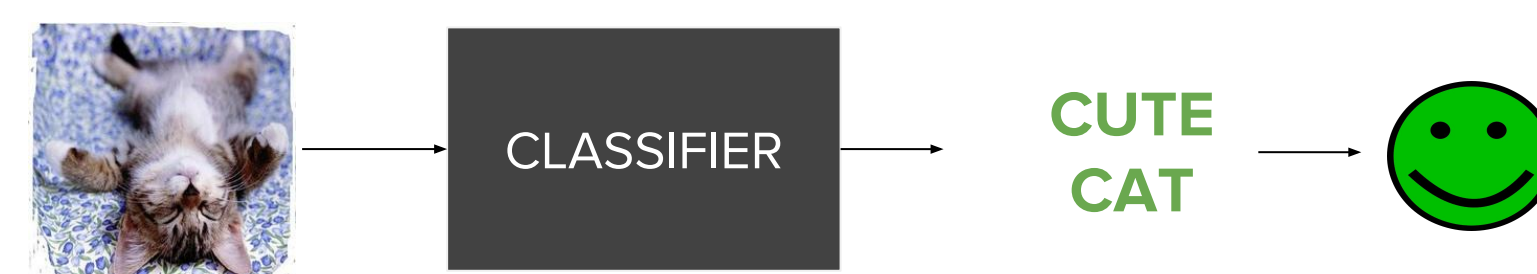
### 1. Goal: Predict Sentiment/Emotions

**Problem:** Affective Gap



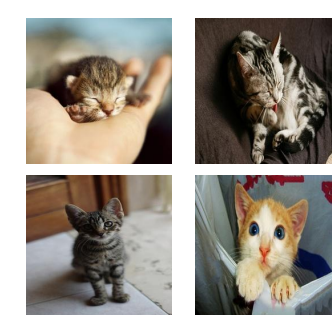
### 2. The Adjective-Noun Pairs (ANPs)

**Solution:** Mid-level Affective Representation



### 3. Hypothesis

Adjective and Noun contribute differently depending on the ANP.

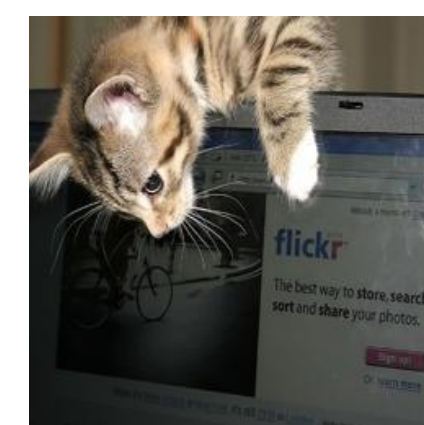


Cute cat



Dark night

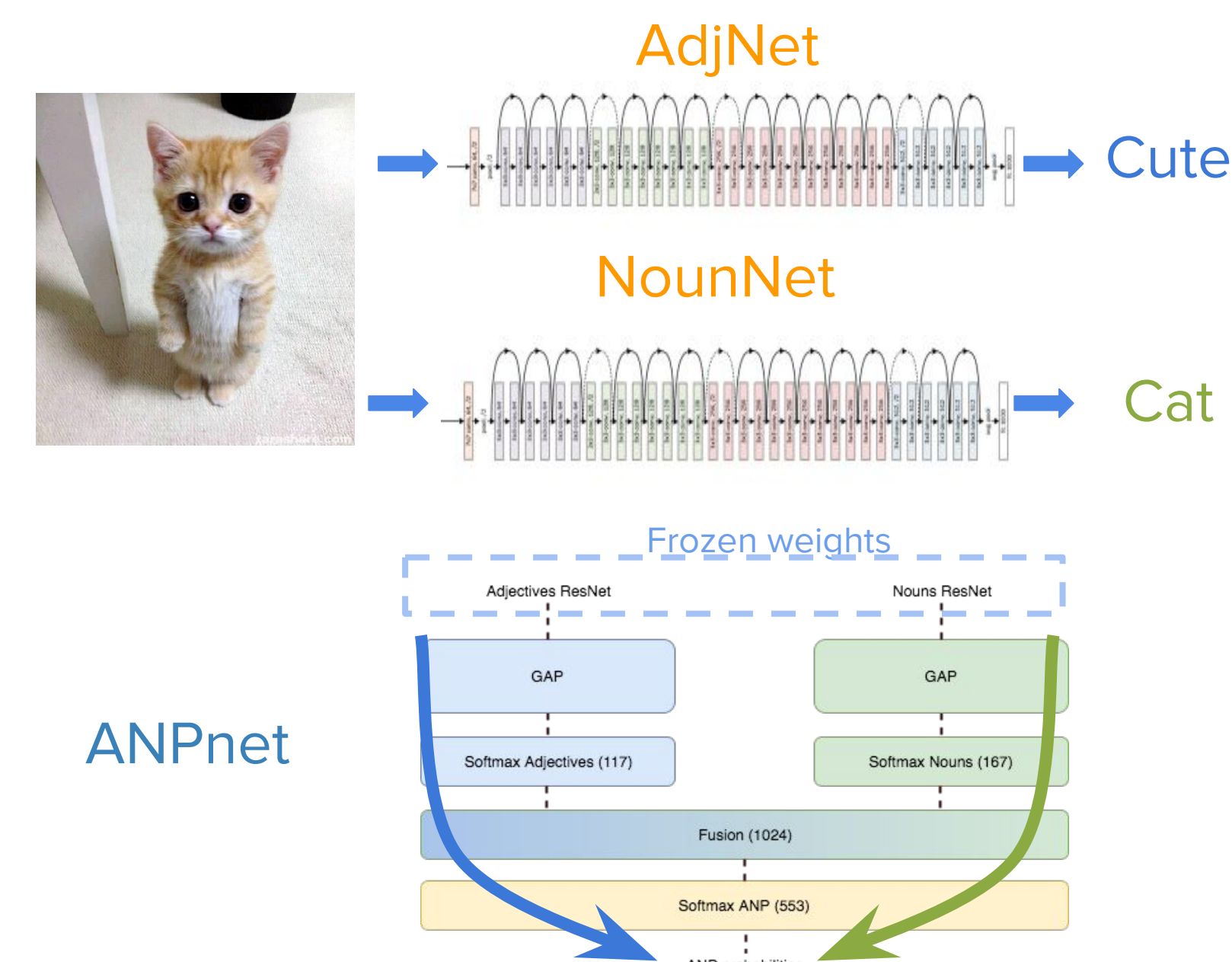
### 4. Dataset (VSO) [1]



Subset of VSO dataset:

- Tag-restricted pool of ANPs [2]
- 553 ANPs with more images
- 384,258 images

## An Interpretable CNN Architecture



Loss in accuracy when forcing an architecture for interpretable predictions

Model	Task	Classes	top-1	top-5
AdjNet [14]	Adj	117	28.45	57.87
AdjNet	Adj	117	27.70	57.00
NounNet [14]	Noun	167	41.64	69.81
NounNet	Noun	167	41.50	69.20
ResNet-50 [14]	ANP	553	22.68	47.82
ResNet-50	ANP	553	23.40	48.20
Non-Interpretable	ANP	553	21.80	46.00
ANPnet	ANP	553	20.67	43.28

## Adjective-to-Noun Ratio

Backpropagating contributions with Deep Taylor Decomposition [3]



	Adjective-to-Noun Ratio (ANR)				All top-5 predictions ANP	Top-5 Accuracy		
	ANP	ANP + Adj	ANP + Noun	ANP + Adj + Noun		Adj	Noun	ANP
sexy model	1.161	1.162	1.163	1.163	1.122	76.52	62.77	59.63
misty trees	1.139	1.140	1.138	1.139	1.146	79.42	71.74	71.88
abandoned places	1.121	1.121	1.033	1.033	1.018	84.58	33.93	8.21
sexy body	1.118	1.118	1.117	1.117	1.110	76.52	57.89	56.44
wild horse	1.117	1.117	1.116	1.117	1.109	54.04	88.50	58.06
innocent eyes	0.787	0.788	0.787	0.788	0.788	43.23	76.44	16.07
incredible view	0.785	0.786	0.785	0.786	0.809	30.71	67.86	39.02
tired eyes	0.776	0.778	0.776	0.788	0.784	56.13	76.44	37.50
laughing baby	0.769	0.769	0.769	0.773	0.773	72.57	83.74	69.03
chubby baby	0.764	0.764	0.764	0.764	0.786	48.00	83.74	45.60

## Equivalent ANPs

Top-5 contributing adjectives and nouns are identical.

Golden Autumn		Golden Leaves	
top-5 adjectives	top-5 nouns	top-5 adjectives	top-5 nouns
golden	autumn	golden	leaves
sunny	leaves	sunny	autumn
colorful	trees	falling	trees
falling	sunlight	colorful	sunlight
bright	tree	bright	tree

## Related ANPs

ANP detected tags can be automatically extended with the most contributing adjective and nouns.

a) Elegant Wedding		b) Charming Places	
top-5 adjectives	top-5 nouns	top-5 adjectives	top-5 nouns
elegant	wedding	charming	hotel
outdoor	cake	comfortable	places
fresh	rose	excellent	house
tasty	dress	traditional	home
delicious	lady	expensive	food



Model and source code:  
<http://bit.ly/musa2>



References:

- [1] Damian Borth et al. "Large-scale visual sentiment ontology and detectors using adjective noun pairs". ACM MM 2013
- [2] Brendan Jou and Shih-Fu Chang. "Deep Cross Residual Learning for Multi-task Visual Recognition". ACM MM 2016
- [3] Montavon, Grégoire, et al. "Explaining nonlinear classification decisions with deep taylor decomposition". ICML 2015



Generalitat  
de Catalunya



GOBIERNO  
DE ESPAÑA  
MINISTERIO  
DE ECONOMÍA, INDUSTRIA  
Y COMPETITIVIDAD

