

Supplementary Material: Injecting Prior Knowledge into Image Caption Generation

Anonymous ECCV submission

Paper ID 8

Image	Captions	Ground Truths
	<p>AoANet: a woman sitting at a table with a cup of coffee.</p> <p>+CLTA+SAE-Reg: a woman sitting at a table eating a hot dog.</p> <p>Top-20 Topic Words: woman, table, holding, eating, girl, sitting, standing, group, hand, food, sandwich, plate, young, front, little, picture, lady, hold, next, man.</p>	<p>GT1: a girl eating a hotdog at a wooden table.</p> <p>GT2: an asian child eating a hot dog sitting at a table.</p> <p>GT3: a little girl posing for a picture while eating food.</p> <p>GT4: a young girl smiles while enjoying her meal.</p> <p>GT5: little girl smiles for the camera as she eats her sandwich.</p>
	<p>AoANet: a little boy sitting on a chair with a laptop</p> <p>+CLTA+SAE-Reg: a baby sitting on a suitcase on the floor.</p> <p>Top-20 Topic Words: luggage, boy, suitcase, bag, child, young, sitting, shoe, elephant, kid, next, little, small, piece, bench, clothes, floor, airport, case, standing.</p>	<p>GT1: young boy sitting on top of a briefcase.</p> <p>GT2: a little boy sitting on a suitcase on the floor.</p> <p>GT3: a toddler boy is sitting on a brief case.</p> <p>GT4: a young baby sits on top of a briefcase.</p> <p>GT5: a small child sitting on top of a briefcase.</p>
	<p>AoANet: a row of motorcycles parked next to each other.</p> <p>+CLTA+SAE-Reg: a row of motorcycles parked on the side of a street.</p> <p>Top-20 Topic Words: wall, street, motorcycle, brick, city, building, row, meter, group, next, parking, side, clock, two, lined, front, sidewalk, parked, road, graffiti.</p>	<p>GT1: a row of motorcycles parked in front of a building.</p> <p>GT2: a number of motorbikes parked on an alley.</p> <p>GT3: a bunch of motorcycles parked along the side of the street.</p> <p>GT4: a back ally neighborhood with motor bikes in a row.</p> <p>GT5: a bunch of motorcycles parked on the side of the road.</p>
	<p>AoANet: a vase filled with white and white flowers.</p> <p>+CLTA+SAE-Reg: a glass vase with some flowers and a doll.</p> <p>Top-20 Topic Words: flower, vase, purple, group, vas, white, table, sitting, filled, glass, rose, watch, watching, together, around, small, large, yellow, next, green.</p>	<p>GT1: a doll standing next to a vase filled with flowers and plants.</p> <p>GT2: a clear vase contains some white pretty flowers.</p> <p>GT3: white flowers in a vase with arranged leaves</p> <p>GT4: flowers are placed in a vase with large leaves.</p> <p>GT5: a flower arrangement in a clear glass vase next to a doll.</p>
	<p>AoANet: a person on a surfboard in the water.</p> <p>+CLTA+SAE-Reg: a man is para sailing in the ocean.</p> <p>Top-20 Topic Words: kite, man, surfboard, water, flying, wave, board, person, ocean, riding, surfer, surf, surfing, group, fly, beach, sky, large, top, red.</p>	<p>GT1: a surfer wrangles a parachute with a scenic mountain background.</p> <p>GT2: a paraglider who has just landed in the ocean</p> <p>GT3: a person riding a surf board with a parachute in a body of water</p> <p>GT4: a man in the water kite surfing on a board.</p> <p>GT5: a man kite boarding over a large body of water.</p>
	<p>AoANet: a display case filled with lots of different flavored donuts.</p> <p>+CLTA+SAE-Reg: a box of donuts with different types of sprinkles.</p> <p>Top-20 Topic Words: group, many, different, several, various, two, together, type, colorful, donut, bunch, watching, around, color, plate, crowd, table, filled, others, gathered.</p>	<p>GT1: a display of yummy looking chocolate covered donuts with sprinkles.</p> <p>GT2: donuts on a tray in a display case</p> <p>GT3: a bunch of doughnuts with sprinkles on them</p> <p>GT4: tray of chocolate donuts with sprinkles in a display case.</p> <p>GT5: a tray full of sprinkle covered, chocolate glazed doughnuts.</p>
	<p>AoANet: a couple of people sitting on a bench.</p> <p>+CLTA+SAE-Reg: two people sitting on a bench looking out at the ocean</p> <p>Top-20 Topic Words: people, two, group, sitting, bench, couple, standing, wooden, park, woman, one, water, next, around, near, several, another, together, sits, crowd.</p>	<p>GT1: a couple is sitting on a bench in front of the water.</p> <p>GT2: two people sitting on a bench silhouetted against the sea.</p> <p>GT3: two people are sitting on a bench together in front of water.</p> <p>GT4: the silhouette of two people sitting on a bench in front of the water.</p> <p>GT5: a couple sits on a park bench and watches the water</p>

Fig. 1: Qualitative examples of generated captions from the baseline AoANet model and AoANet with our proposed CLTA and SAE Regularizer. We also visualize the Top-20 Topic words from the learnt latent topic space in our CLTA module.