**SITUAÇÃO 1**

**Situação:** as mesmas entidades e a mesma label, mas relação diferente (diferente quer dizer oposta)

**O que contrastar:**

* Imagem (captar a relação especial diferente na imagem, attention a partes das imagens deve ser diferente e por isso embeddings devem ser diferentes para relações diferentes porque a informação da relação diferente está contida na imagem)
* Texto (dar significado diferente à palavra da relação, que é oposta, leva a embeddings diferentes)
* Imagem-texto, texto-imagem (??? relação entre a palavra e a relação especial na imagem e vice-versa – image-to-text attention - a relação é a oposta por isso os embeddings visuais e textuais devem ser mais similares entre a palavra e a imagem quando a relação é diferente e devem ser mais diferentes quando a relação é oposta, partindo do pressuposto que estes embeddings se encontram no mesmo espaço. Para isto é obrigatório que os embeddings sejam primeiro projetados com uma layer de projeção???)

**Como contrastar:** Inverter as labels de uma relação do par de relações (ou as negativas ou as positivas), uma vez que se não inverter isso induziria um erro, uma vez que a relação não é a mesma mas a oposta

**Dúvidas:** Posso contrastar só com samples da mesma relation e com label sempre negative/positive?

**Pares de relações consideradas opostas / Exemplos :**

* group1 = ['on','on top of','under','at']
* group2 = ['behind','in front of','at the back of']

label 0

A group of people sitting under an umbrella

Description automatically generated with medium confidencevs A horse looking out a window

Description automatically generated with low confidencevs A picture containing tree, outdoor, mammal, horse

Description automatically generated

* group3 = ['at the left side of','at the right side of','next to','left of','right of','at the side of','congruent']

A picture containing text, indoor, food, bread

Description automatically generated vs A plate of food next to a computer

Description automatically generated with medium confidence

* group4 = ['contains','in','inside','off','within','outside','into','out of']
* group5 = ['beneath','above','below','over','down from']
* group6 = ['at the edge of','far away from','beside','far from','away from','near','close to','adjacent to','attached to','by']

Graphical user interface

Description automatically generatedvs A picture containing text, horse, transport, mammal

Description automatically generated

* group7 = ['facing away from','facing']

Label 0

A picture containing text

Description automatically generatedvs Graphical user interface

Description automatically generated with low confidence

* group8 = ['part of']
* group9 = ['surrounding','around','among','enclosed by']
* group10 = ['across from','opposite to','across']
* group11 = ['parallel to','alongside','perpendicular to','along']
* group12 = ['in the middle of','between']
* group13 = ['against','toward']
* group14 =['ahead of']
* group15 = ['connected to','detached from','touching']
* group16 = ['has a part']
* group17 = ['consists of']
* group18 = ['with']
* group19 = ['beyond','past']

**SITUAÇÃO 2**

**Situação:** a mesma relação e as mesmas entidades mas com uma label diferente

**O que contrastar:**

* Imagem (faz sentido contrastar porque a relação especial é diferente nas imagens, a attention deve captar relações diferentes através de pestrar atenções a partes (relações) diferentes da imagem e consequente os embeddings também serão diferentes)
* Texto (não faz sentido porque é igual a caption, apesar de a label ser igual)
* Imagem-texto, texto-imagem (??? não faz sentido contrastar porque não queremos que os embeddings sejam contrastado com o texto porque o mesmo é sempre igual mesmo com label diferente)

**Como contrastar:** Não é preciso inverter as labels porque a relação é a mesma e não oposta

**Exemplos:**