**Cross-modal Search for Fashion Attributes Summary**

**Source:**

**https://kddfashion2017.mybluemix.net/final\_submissions/ML4Fashion\_paper\_7.pdf**

Goal: to develop a neural network model by feeding:

1. Incomplete image descriptions
2. But clear images

Tasks to be carried out:

* **Segment the images and product descriptions.**
* Proposed a rule based image segmentation approach to perform it.
* An objective function is designed that encourages the model that can bring up some kind of similarity between one text fragment and image fragment(i.e. they have a high inner product).
* To simplify, we can say that similar looking images should have same descriptions.
* Images with missing or noisy descriptions can be hence recovered with the help of objective function.
* The paper proposes a **cross model** search tool that can learn from noisy data.
* **Cross model means given an image, return suitable textual description or given a certain text retrieve related images.**

Steps followed:

1. First selective search segmentation : to show regions of the image that contain fashion attributes
2. Rule based segmentation: Applying certain rules to detect the image from background and face.
3. Image representation: Once image fragments are obtained after step 2, a CNN network is trained on ImageNet.
4. Text segmentation: Word embeddings on product descriptions are trained using Skip-gram model, performed NLP.
5. Cross model is trained.