**Image Similarity using Deep Ranking**

* Image similarity is the the measure of how similar two images are.

In other words, it **quantifies the degree** of similarity between **intensity patterns** in two images

\*\* Deep understandings : image에 대한 intensity pattern에 대해 명확히 이해하기 위해서는 Intensity Transformation에 대한 정의로 이어져야 한다.

* 가장 기본적인 정의는 이와 같다.

입력이미지를 구성하는 각 픽셀들의 Inensity value를 수학적 식에 따라 대응 출력하여 이로 해당 이미지의 결과값에 대하여 mapping하는 과정이라고 이해하면 되겠다.

여기서, Transformation에 대한 기준을 무엇으로 기준하느냐에 따라, 해당 결과값은 달라지게 된다. 즉, 이미지에 대한 해석 또는 번역에 대하여, 그 기준을 잡아야 한다는 것이다.

* Negative Transformation
* Log Transformation
* Power-Low(Gamma) Transformation
* Piecewise – Iinear Transformation functions 등이 대중적인 Transformation 이며 해당 Transformation에 대한 깊은 내용은 본 Research review와는 조금 동떨어지므로 추 후 필요시 재거론하도록 하겠다.
* **How to build an image similarity model**

**:** image similarity models considered **category-level image similarity**.

**x(but, This category level image similarity is not sufficient for the search-by-example image search application.)**

: \*\* Search by query image를 가능하게 하기 위해서는 동일 카테고리 내에서도 서로 다른 이미지 간에 차이를 구분할 수 있어야 한다.

* 이를 위한 방법 중 하나로 image similarity model은 features를 추출해낸다.
* Global filters

: it is a linear filter used for texture analysis

* + - * + Analyzes whether there is any specific frequency content in the image in specific directions in a localized region around the point or region of analysis.
        + **They have been found to be particularly appropriate for texture representation and discrimination.**
        + In the spatial domain, a 2-D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave.
* Local Binary Patterns, HOG
* Using these feature to compute the similarity between the features.
* **\*\*\* Deep ranking, which learns fine-grained image similarity by characterizing the fine-grained image similarity relationship with set of triplets.**

Query

Positive

Negative

구분 : Image Similarity / Image Classification

* For a similar image ranking model, it would look at the colours and other aspects of the cars as well.
  + - * + Ex\_ If a query image is a black car, the similar image ranking model would rank the dark gray car higher than than the white car.
* Deep ranking model characterizes the fine gained image similarity relationship with a set of triplets.
* How to :

: Using metric

L1-norm (Manhattan distance)

L2 norm (Euclidean distance)

: The Manhattan distance may better capture the human notions of image similarity.

: In the Deep ranking, Squared Euclidean distance as the similarity metric.

* + - * + The maller the distance the more similar the images are.

텍스트이(가) 표시된 사진

자동 생성된 설명

* \*\* The whole Deep ranking architecture can be thought of as a **function that would map the image to a point in the Euclidean space.**
* The goal is to learn an embedding function that assigns smaller distance to more similar images.

f, embbeding function to would map the image to a vector.

Pi is the query image

Pi+ is the positive image, Pi- is the negative image.

r, similarity distance between two images.

The hinge loss for the triplet is defined as:



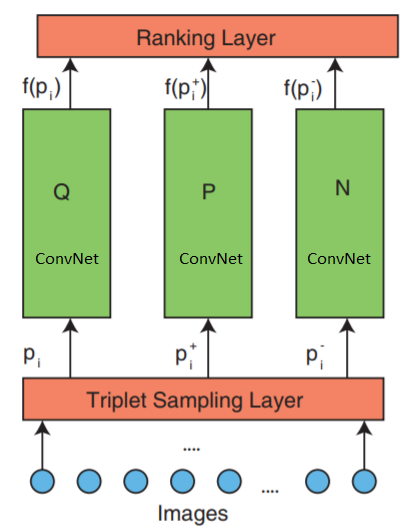
**Where ‘l’ is the hinge loss** for the triplet, g is a gap parameter that regularizes the **gap between the distance of the two image pairs**: (Pi, Pi+), (Pi, Pi-)

D, Euclidean distance between the two euclidean points.

Optimising the model in such a way that,

The distance between query image and positive image is not only lesser than the distance between query image and negative, **but is lesser by an ‘g’**

**-Network Archtecture of Deep Ranking-**

**3parts**

- **Triple sampling**

* **ConvNet**

**- Image Similarity ranking**

**: One image triplet contains a query image pi, and Pi+, Pi-**

**: Which are fed independently into three identical deep neural networks ‘f(.)’ with shared architecture and parameters**

The triplet characterizes **the relative similarity relationship** for the **three images**.

\*\*\* **The deep neural network ‘f(.)’ computes the embedding of an image ‘pi’ : f(pi)** ∈ Rd

Where ‘d’is the dimension of the feature embedding, and ‘R’ represents the real number space.