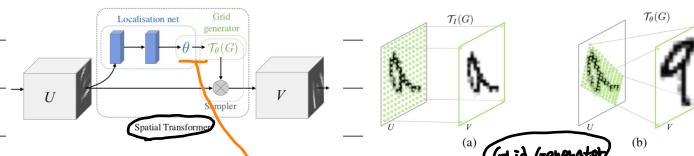


**STN** 개념이 쉬워서 할게 없다.

- CNN은 강력하지만 spatially invariant 하기 때문에 제한적인 부분이 존재한다.
- 네트워크에서 공간적으로 데이터를 조작할 수 있는 Spatial Transformer이다.
- 이 모듈은 최적화 과정에서 여분의 학습 조정이 필요없이 조건 하에서 스스로 특징맵을 공간적으로 변환할 수 있다.

### 요약

- CNN의 한계 중 하나는 공간 상에서 크기가 변하거나 회전되거나 위치를 이동하였을 때 즉, 공간적인 변화가 있을 경우 문제를 잘 탐지하지 못하는 문제가 있다.
- Spatial transform은 저그러트거나 회전시킨 등의 노이즈가 첨가된 이미지를 통해 적절히 복원하여 적절한 아웃풋을 도출해 내는 모듈이다.
- Localisation net은 input을 통해 적절히 복원할 output을 만들어낼 transform matrix를 찾는다.
- Grid Generator은 transform matrix (localisation set에서 찾아진)를 실제적으로 매핑하는 단계이다.
- Sampler는 Grid Generator가 만들어낸 값들을 하나하나 읽어오게 된다.



이렇게 얻은  $\theta$ 는 아래와 같은 방식으로 병렬된다

$$\begin{pmatrix} x_i^t \\ y_i^t \end{pmatrix} = T_\theta(G_i) = A_\theta \begin{pmatrix} x_i^i \\ y_i^i \\ 1 \end{pmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{bmatrix} \begin{pmatrix} x_i^i \\ y_i^i \\ 1 \end{pmatrix}$$