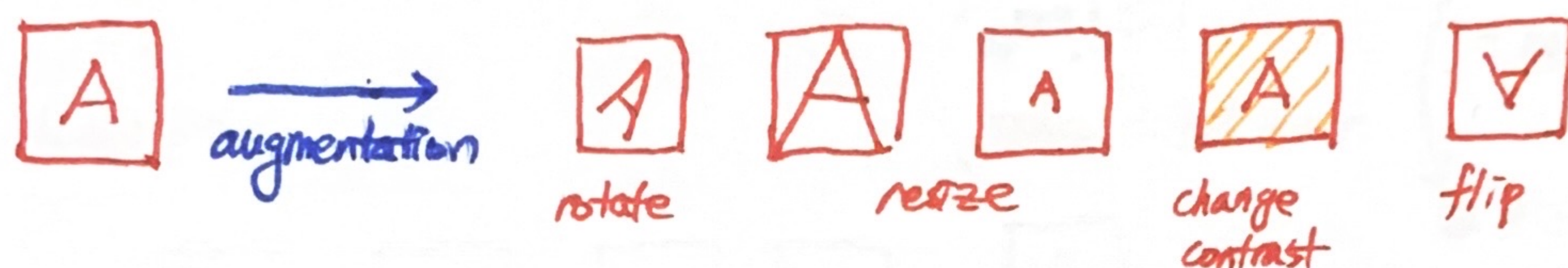


<Machine Learning development process - Data Augmentation>

- Augmentation: modifying an existing training example to create a new training example.
- trying to get more data of all types can be slow and expensive
- alternative way of adding data of the types where error analysts has indicated might help
- another technique widely used for images and audio data that can increase training set size is data augmentation

ex) alphabet letter image



* Data augmentation by introducing distortions

- augmentation does not help to add *purely random/meaningless noise* to your data.

아무 의미도 없고 무작위로 noise를 추가하여 augmentation 하는 것은 모델 학습에 큰 도움 X

⇒ test set에 있는 noise나 distortion 유형들은 나타내는 noise를 도입하여 augmentation 해야됨

<Data Synthesis>

- Synthesis: using artificial data inputs to create a new training example.

↳ not by modifying an existing example but by creating brand new example