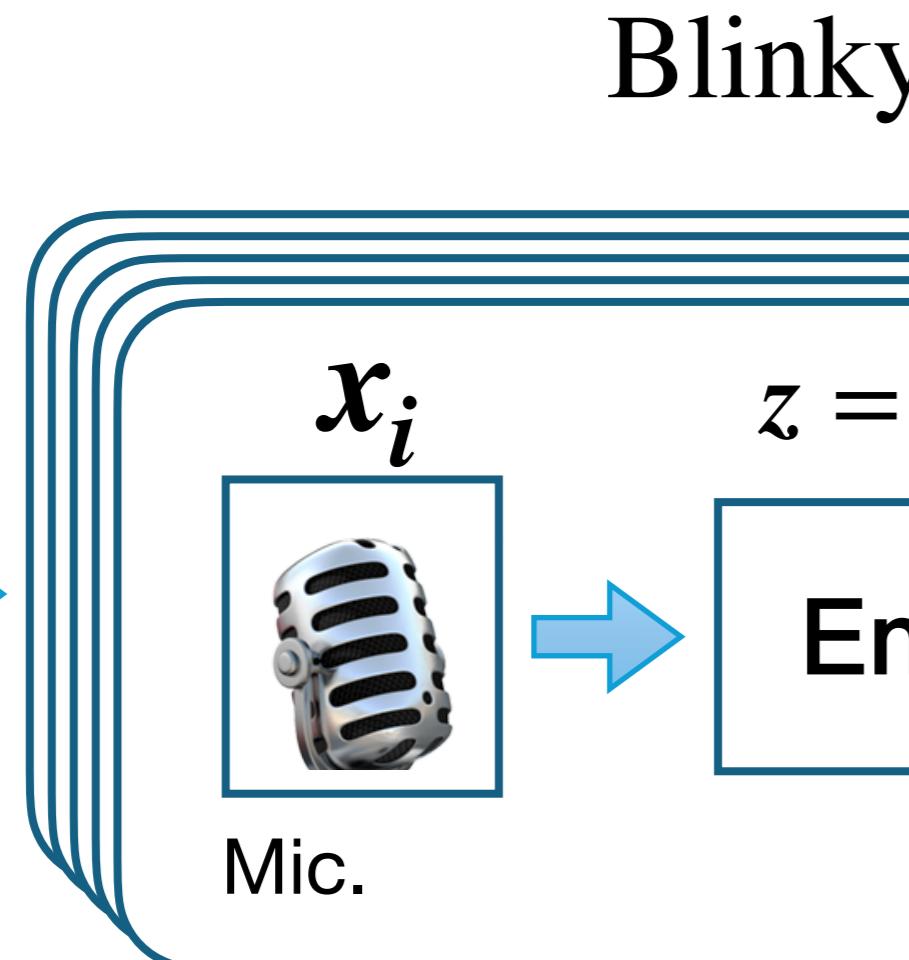
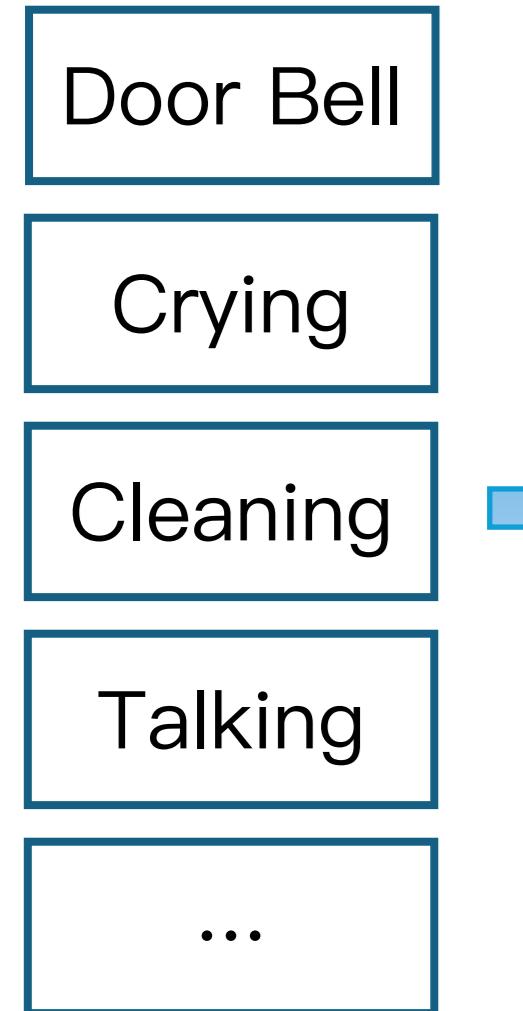


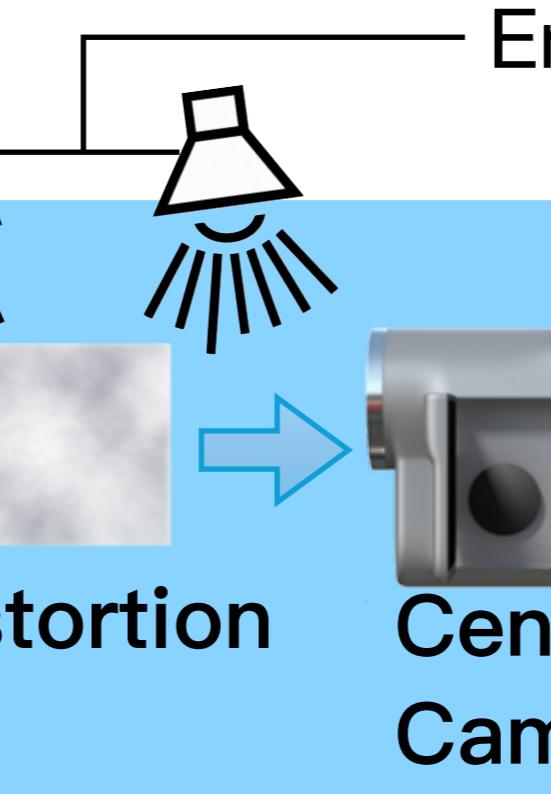
Acoustic Event



$z = \text{Embed}(x_i)$

Embedding

$z'_i = \text{Transmit}(z_i)$



$z'_i = \text{Transmit}(z_i)$

Environment Light noise

Central Camera

Temporal LED Intensity Variation

Classifier

One-hot Prediction

\hat{y}