Co-Clustering to Reveal Salient Facial Features for Expression Recognition

- **Proposal** Use co-clustering to select features that can be used to classify facial expressions
- **Method**
  - Use Gabor filter to extract features from facial images
  - Use co-clustering to attain a subset of features and samples
  - Find the probability that a co-cluster is related to a certain class
  - Features with high probability are selected

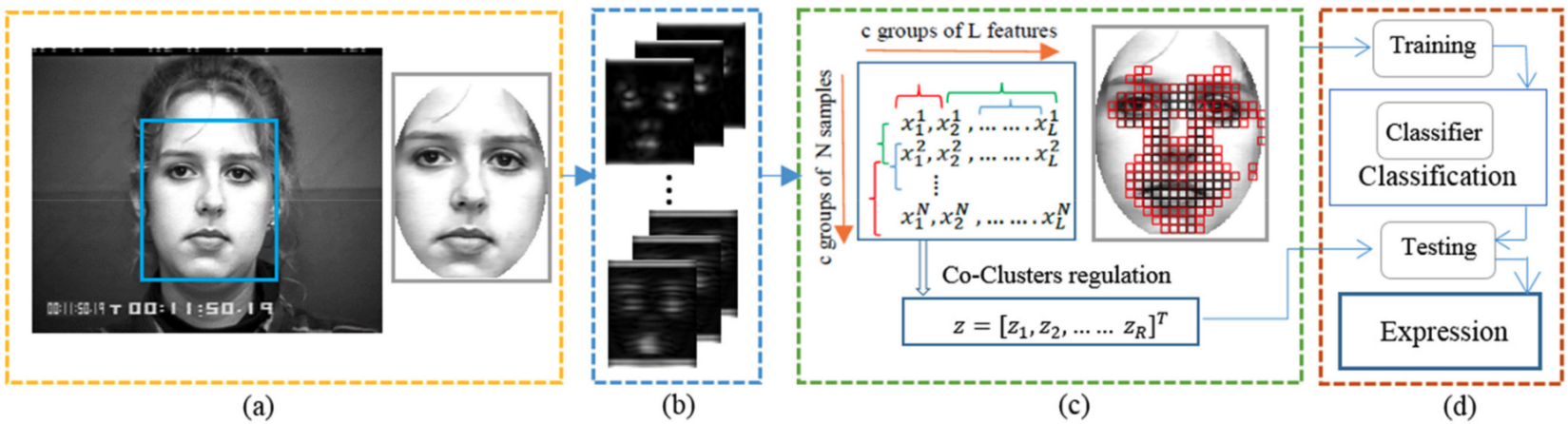

Figure: Khan et al. 2017

My Proposal

- Word-embedding or n-gram to extract features from text
- (Maybe) Generate more features from the results above
- Use co-clustering to attain a subset of features and samples
- Select features that are related to a certain class most

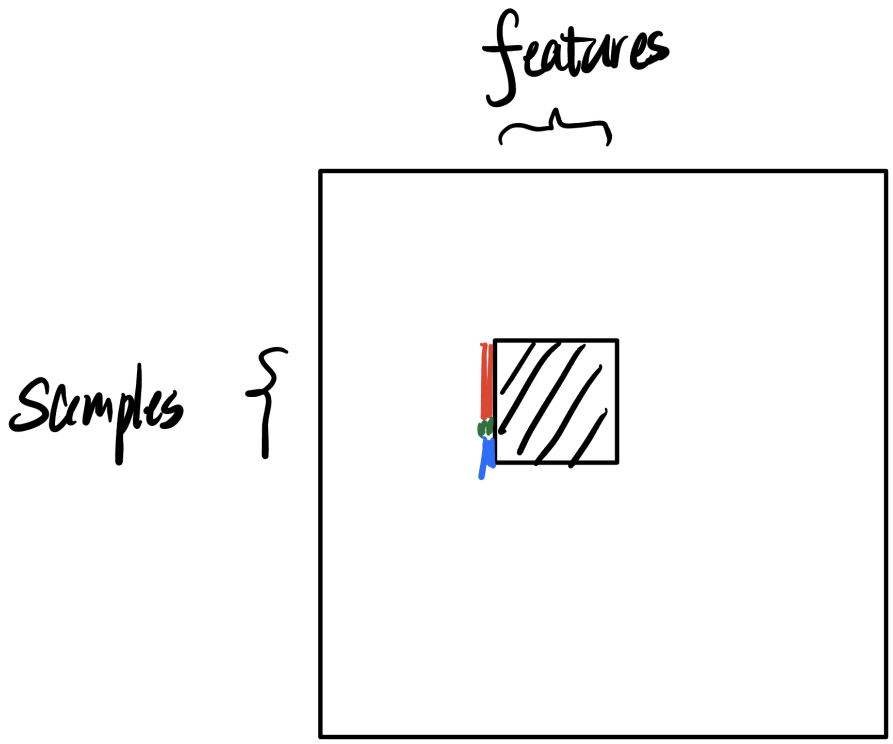

Figure: My Proposal, colors for classes

Understanding Black-box Predictions via Influence Functions (ICML 2017 Best Paper)

- **Introduction** This paper uses influence function to measure the importance of each training sample and feature. It can be applied to any model that is twice differentiable.

$$z_\delta \overset{\text{def}}{=} (x+\delta, y), \hat{\theta}_{\epsilon,z} \overset{\text{def}}{=} \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{i=1}^{n} L\left(z_i,\theta\right) + \epsilon L(z,\theta)$$

$$\hat{\theta}_{\epsilon,z_\delta,-z} \overset{\text{def}}{=} \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{i=1}^{n} L\left(z_i,\theta\right) + \epsilon L\left(z_\delta,\theta\right) - \epsilon L(z,\theta)$$

$$\mathcal{I}_{\text{up,params}}\left(z\right) \overset{\text{def}}{=} \left.\frac{d\hat{\theta}_{\epsilon,z}}{d\epsilon}\right|_{\epsilon=0} = -H_{\hat{\theta}}^{-1}\nabla_\theta L(z,\hat{\theta})$$

$$\mathcal{I}_{\text{pert,loss}}\left(z, z_{\text{test}}\right)^\top \overset{\text{def}}{=} \left.\nabla_\delta L\left(z_{\text{test}}, \hat{\theta}_{z_\delta,-z}\right)^\top\right|_{\delta=0}$$

$$= -\nabla_\theta L\left(z_{\text{test}}, \hat{\theta}\right)^\top H_{\hat{\theta}}^{-1}\nabla_x\nabla_\theta L(z,\hat{\theta})$$