平成 32 年 9 月 11 日　　細川　岳大

# 進捗報告

## 1　今週やったこと

- Virtual Adversarial Training(:VAT) の実装中

## 2　Virtual Adversarial Training(:VAT)

Virtual Adversarial Training(:VAT) は事後確率の分布を滑らかにして汎化性能の向上させることを目的とした手法である．

### 2.1　損失

事後確率の分布を滑らかにするためにデータ $\boldsymbol{x}$ 自身とそれに小さな摂動 $\boldsymbol{r}$ を付与したもの事後確率の距離を近づけることでそれを達成する．$KLD\left[p(\boldsymbol{y}|\boldsymbol{x}), p(\boldsymbol{y}|\boldsymbol{x}+\boldsymbol{r})\right]$ を最小化するためこれの平均値を損失として加える．
また，$\boldsymbol{r}$ について全方向で行うのは現実的でないため $KLD\left[p(\boldsymbol{y}|\boldsymbol{x}), p(\boldsymbol{y}|\boldsymbol{x}+\boldsymbol{r})\right]$ を最大化する $\boldsymbol{r}_{vadv}$ についてのみ算出する．

前述の $KLD$ を $D(\boldsymbol{r}, \boldsymbol{x})$ とすると $\mathbf{r}_{vadv}$ を求めるために，$\boldsymbol{r} = 0$ でテイラー展開する

$$D(\boldsymbol{r}, \boldsymbol{x}) \approx D(\mathbf{0}, \boldsymbol{x}) + \boldsymbol{r}^{\mathrm{T}} \nabla_{\boldsymbol{r}} D(\boldsymbol{r}, \boldsymbol{x})|_{\boldsymbol{r}=\mathbf{0}} + \frac{1}{2}\boldsymbol{r}^{\mathrm{T}} H(\boldsymbol{x})\boldsymbol{r} \quad (1)$$

$\boldsymbol{r} = \mathbf{0}$ から第三項のみが残る．このとき $\boldsymbol{r}$ は $H$ を最大化する固有ベクトルにあたるので，べき乗法による計算で求める．またそのときの $H$ を差分法から求めらる

### 2.2　実装について

keras を用いて試行錯誤やってみたが自分の力量では実装するまでに至らなかった．
pytorch での実装がすでになされていたので，今週はそれを動かしつつ pytorch の勉強中．

## 3　来週の課題

- SSL の実装及び実験