

置き換え対象分類器との分類結果の互換性を維持するための、“Variational AutoEncoder”を利用した類似データ生成手法および部分的再学習手法の提案

富士通株式会社 吉田裕輔 yoshida_yusuke@jp.fujitsu.com

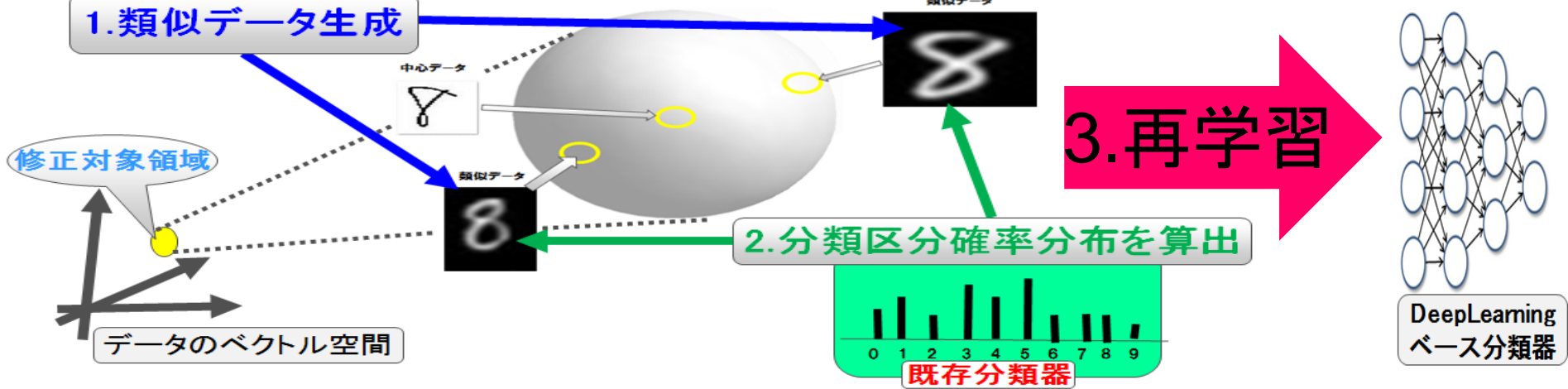
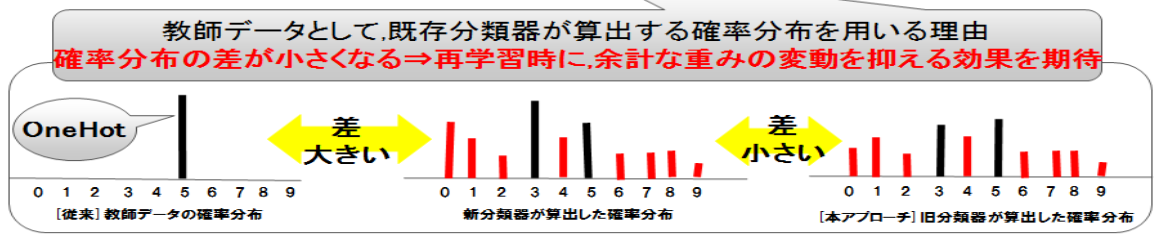
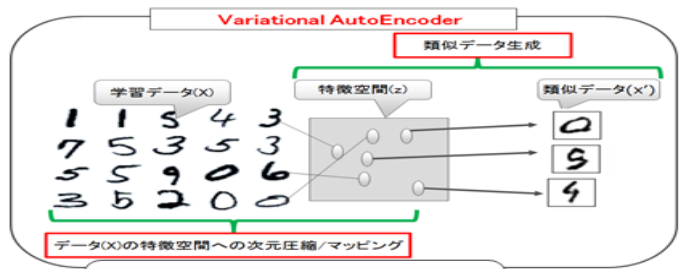
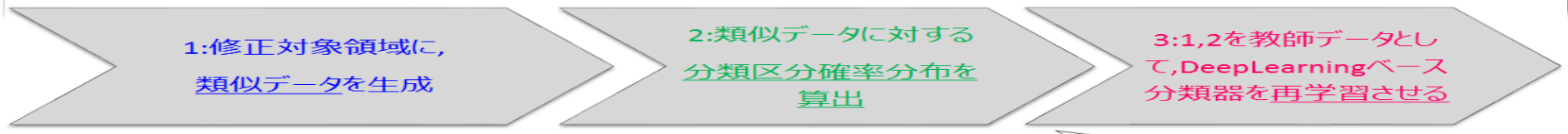
背景と問題

既存分類器をDeep Learningベース分類器に置き換える際、後者が部分的に能力が劣る場合があるが、従来手法(モデル/パラメータチューニング後の、学習のやり直し)では、その能力差のみを解消することは極めて難しい。その結果、置き換え後にレベルダウンのリスクが残る。

解決のアイデア=部分的再学習

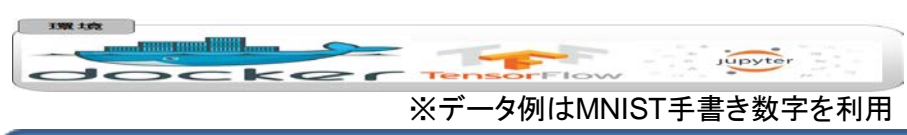
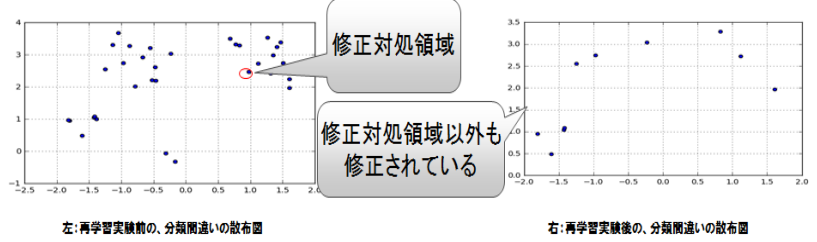
- ◆ 分類境界の一部分のみを、既存分類器の分類境界に合わせる
- ◆ 修正対象領域以外のデータに対する分類結果を変更しない

アプローチ



実験結果

- ◆ 修正対象領域のデータに対しては、正解するようになった
- ◆ 修正対象領域以外のデータに対しては、分類結果が変わった



今後の課題

- 本アプローチでの推測
OneHot分布との差より、新分類器の確率分布と旧分類器の確率分布との差の方が小さく、そのため損失関数E(=クロスエントロピー)が、小さくなると推測していた。
- 分析: 影響を抑えたい損失関数Eの大きさと、それによる各ユニットの重みの変動を比較
各出力層ユニットにおける損失関数は、具体的には以下の式である。
$$E_k = -\{t_k \log(y_k) + (1 - t_k) \log(1 - y_k)\}$$

k: 確率変数 0~9, y_k : 確率変数kにおける、新分類器での確率値
 t_k : [従来] OneHot分布(1 or 0)であり、k=0~9で正解の数字の場合のみ“1”で、他は“0”
[本アプローチ] 旧分類器が算出した確率分布における確率値であり、0~1の値
正解ではない数字における損失関数Eの影響を抑えたいと期待していたため、以下2つを比較する。
◆ [従来のE]: $-\log(1 - y_k)$
◆ [本アプローチのE]: $-\{(0 \sim 1) * \log(y_k) + (1 \sim 0) * \log(1 - y_k)\}$
次に、上記の損失関数値から、各ユニットの重み ω の変動 $\Delta\omega$ を比較する。

