大学共同利用機関法人 情報・システム研究機構
国立情報学研究所
National Institute of Informatics

トップエスイー ソフトウェア開発実践演習

人工知能とテスト検証 Test4AI

石野 克徳, 上野 一輝, 岡本 将明, 小澤 遼, 小山 尚晃

DL技術のテスト手法の問題

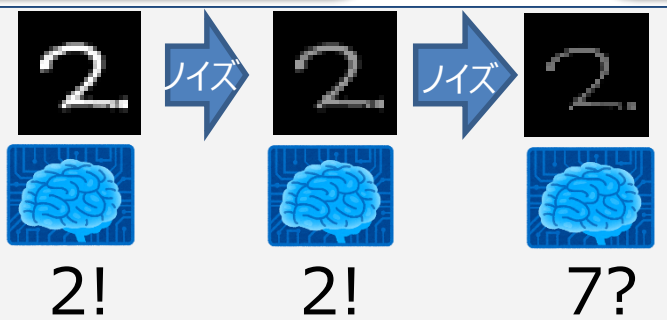
DL (深層学習) 技術が社会で広く適用されている中、自動運転など安全が重要な領域を中心に DL 技術のテストの重要性が高まっている.
しかし、学術的に研究はされているものの、産業での利用にはハードルがある.

手法・ツールの選定と調査

DL テスト技術の先駆的手法である DeepXplore を実践するとともに、次の3観点で網羅的な調査を行った。
- パラメタ調査、効率化検討
- 複数ドメインのモデルへの展開
- 内部処理の調査

DeepXplore手法

発火していないニューロンを発火させるように入力画像に自動的にノイズを加える



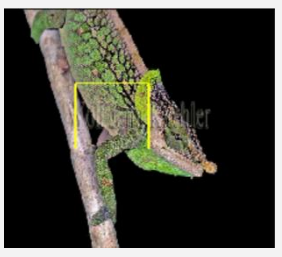
ノイズをかけ続け，誤認識する画像を生成！

パラメータ(設定値)
- ノイズの種類 (occl,blackout,light)
- ノイズをかける位置
- 最急降下法ステップサイズ
- ノイズをかける回数　など

1. パラメタ調査・効率化検討

DeepXplore はノイズの位置をパラメータとして与える必要がある
モデル解釈手法 LIME を利用し,特徴的な効率化領域をノイズ位置に自動設定した.

LIME出力



DeepXplore入力

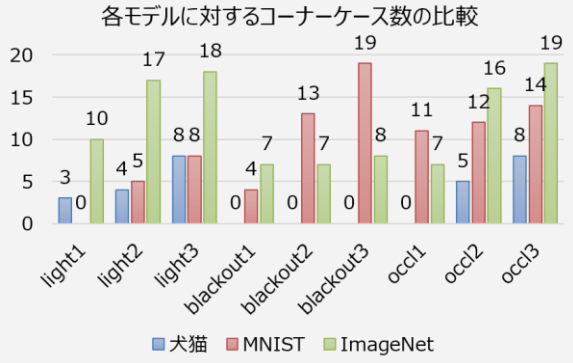
パラメータ固定と自動設定を比較画像50件に各10回適用し出力件数を検証

	自動	固定
合計	180	166
平均	18	17
中央値	18	17
不偏分散	2.9	7.2

自動設定により**効率化**と
安定した結果の獲得に成功

2. 複数モデルへの展開

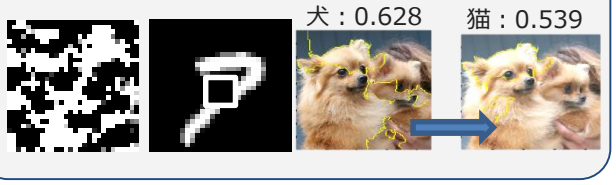
犬猫,MNIST,ImageNetの3つのモデルに対し,様々なパラメータ設定でDeepXploreを適用し,誤識別画像の生成数の比較と考察を行った.



①強いノイズ条件よって検出率は増加するが,極端なノイズの画像が増加

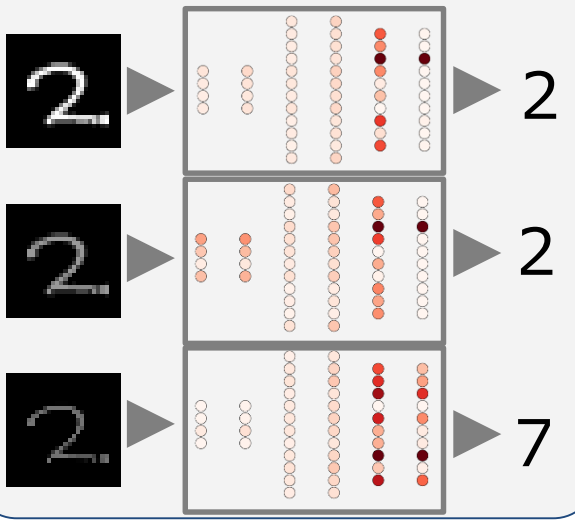
②MNISTに対してblackoutが問題設定に合っており,高い検出率を示した

③元画像の確度が検出率に影響するため、ImageNetでの検出率が高い

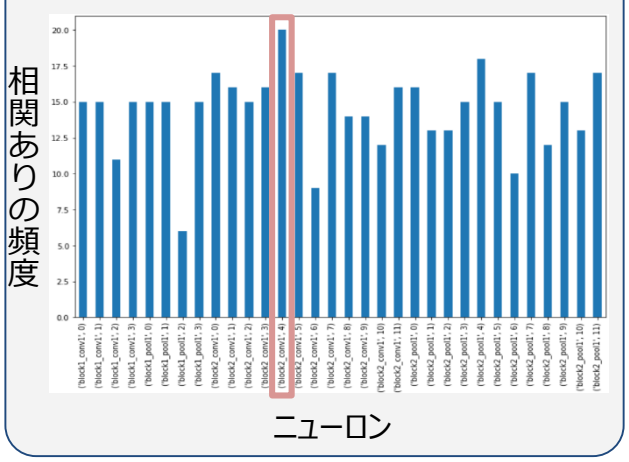


3. 内部処理の調査

ノイズによるニューロン発火の変化



誤識別と相関するニューロン特定



トップエスイー　サイエンスによる知的ものづくり教育プログラム