

# 学習レス型確率分布データセットモデルを用いた 欠陥検出，類似画像探索，及びクラスタリング

永田研究室 F122065 山口 桐

## 1. 目的

近年高齢化社会となり熟練した観察技術を持つ品質検査の担当者から経験の浅い担当者への変更を余儀なくされることが少なくなく，生産ラインにおける品質検査能力の低下が課題となっている。そこで過去に発生した不良品の製品画像データをもとに新たに発生した欠陥との類似度を定量的に算出することで，欠陥検出の精度を向上させる方法について検討する。また，提案手法を用いて葉の視覚的な特徴の違いや変化を定量的に扱うことで，作物の適切な収穫時期や病気への遷移状態の予測など，農業分野への応用についても検討する。

## 2. 研究内容

本研究では既に公開されている畳み込みニューラルネットワーク(CNN)モデルを特徴抽出器として一定量の画像データから抽出した多次元の特徴ベクトルを確率ベクトルに変換後データセットとして構築しておくことで類似画像の探索に用いる。確率ベクトルとのカルバック・ライブラー情報量(Kullback-Leibler Divergence : KLD)を計算し，その大きさを評価することで画像分類やクラスタリングを可能としている。学習レス型確率分布データセットを用いた分類は，従来の CNN による分類のように入力した画像を分類するのではなく CNN の全結合層部を特徴ベクトル抽出に利用し，多次元(例えば，4096 次元)の特徴ベクトルを確率ベクトルに変換後，データセットとして構築しておく。その後，構築したデータセット内の全ての確率ベクトルとテスト画像の確率ベクトルとの KLD 値を計算し，その値の最も小さかったクラスにテスト画像は分類できる。実験では，広く普及している転移学習で構築された CNN と，提案する学習レス型確率分布データセットモデルとの分類精度の比較を行う。また，提案する学習レス型確率分布データセットモデルが有するクラスタリング機能は，ある未知の画像の特徴確率ベクトルと，クラスタリングしたいデータセット内の全ての特徴確率ベクトルとの KLD を計算し，その大きさをもとにグループ分けを行う。このクラスタリング機能を用いることで，健康な状態から病気へと遷移する途中の過渡的な症状も判別できるようになるものと期待される。

## 3. 結果

VGG19 の転移学習で構築した CNN (分類法 1) と，それを特徴抽出に用いた学習レス型確率分布データセット(分類法 2) の分類精度を比較するために，それぞれ 5 回の試行を行った結果を表 1 に示す。なお，分類法 2 では分類法 1 の CNN の全結合層部第 39 層にある fc6 層から得られる特徴ベクトルを用いている。表 1 の結果から 2 つの分類法においては分類精度に大きな差がないことが確認された。また，ある未知の画像に対して類似画像の探索機能を適用すると，画像内の葉の構図に加えて葉の表面に似た症状を持つ画像がデータセットの中から良好に検索されていることが確認された。さらに，図 1 のようにある探索元画像に近い特徴を持つ画像を KLD 値のレンジをもとにグループ化することでクラスタリングも可能となった。クラスタリングを使用することで従来の分類法では見つけることが困難であった「健康な状態から病気へと遷移する状態」の検出も可能になるものと考える。

Table 1 Comparison of Classification.

Trial	Method 1	Method 2
CNN 1	89.8%	89.0%
CNN 2	89.8%	88.7%
CNN 3	88.6%	87.2%
CNN 4	91.8%	91.6%
CNN 5	86.6%	87.2%

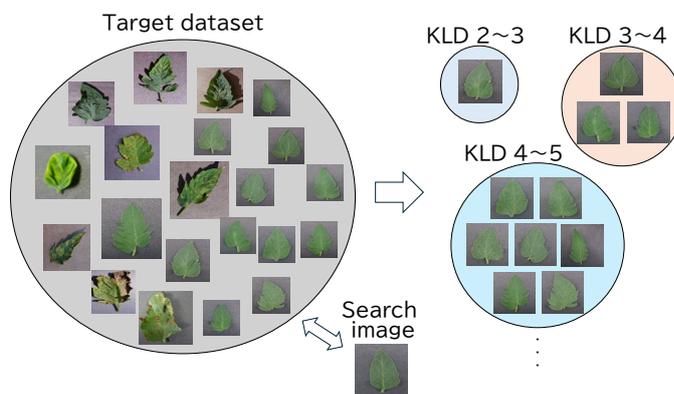


Fig. 1 Example of clustering result.