

## 2章 画像認識を用いた外観検査の開発方法

リコー 笠原 亮介

本章では、近年人工知能(AI)技術として脚光を浴びている機械学習ベースの画像認識技術を外観検査へ応用する場合の考え方、開発方法について解説する。

### 機械学習による外観検査開発フロー

画像を用いた外観検出技術として、欠陥検出技術の開発フローを図1に示す。以下の節でそれぞれ、入力画像データ・特徴量・機械学習・性能評価の方法について述べる。

#### 入力画像データ

ここでは、入力画像データである学習サンプル

やテストサンプルについて述べる。

#### 1. 正確なラベル付け

機械学習を行うため、画像ごとに、写っている対象物は良品、あるいは欠陥品であるというようなラベルデータを付ける必要がある。その際、最終的に高い認識精度を出すためには、学習データの正確なラベル付けが必要となる。通常、ラベル付けは人間が行うため、ラベル付けの際の見逃しなどによる、誤ったラベル付けや、複数人でラベルを付ける場合には、人によって基準がずれることにより、大量のデータに対して、適切なラベルを付けることは思いのほか難しい面がある。そのため、事前に統一した基準を設けて、ラベル付けを行う作業者に十分に周知・すり合わせをしておく必要がある。

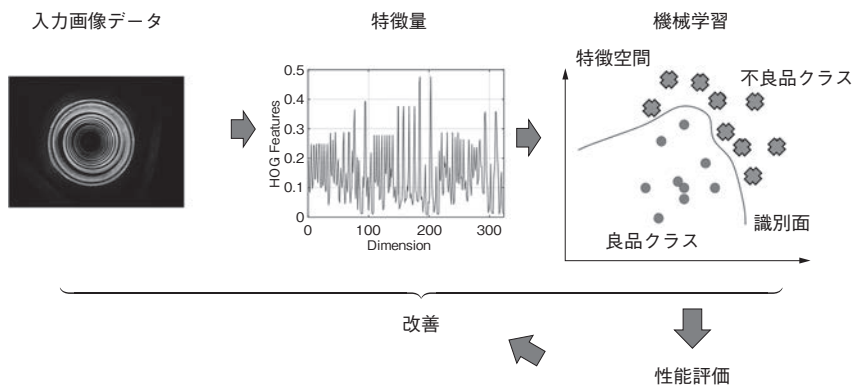


図1 画像を用いた欠陥検査技術の開発フロー

## 2. 大量データの収集

高い精度を出すためには、表現能力の高い、複雑なアルゴリズムを使いたいが、そのためには大量のデータが必要となる。しかし、検査対象によっては大量にデータを集めることが難しい。特に、工程ラインにおいて、欠陥品は良品に比べて絶対量が少ないため、さらにデータ収集が困難である。

必要な学習サンプル数としては、機械学習の観点からは、後述する「過学習」の問題を回避するため、一般に特徴量の次元数の5~10倍のデータが必要である。特徴量の次元数としては、一般的に単純な特徴量だと数千次元から数十次元、複雑な特徴量だと数百から数千次元であるため、データ数が少ない場合には、単純な特徴量を使った方が良い場合も多い。

また、認識精度を評価するために、学習するサンプル以外にテストサンプルが必要となる。このテストサンプルの数は、要求精度の観点から最低限で、例えば未検出率1%を実現するためには、100サンプル、未検出率0.1%を評価するためには、1000サンプルが必要となる。必要な未検出率が小さい場合には、それだけテストサンプルの数も必要となることに注意が必要である。

## 3. Data Augmentation

前述のように、かなり多数のサンプルが必要となるが、サンプル数の取得に時間的・コスト的な制限があり、得られるデータが少ない場合が多い。その場合、Data Augmentationという方法が用いられる。これは、擬似的にデータを増やす手段であり、準備した単一のサンプル画像に対して、クロッピング、回転、ミラー、輝度変更、あるいは空間周波数のフィルタリングなどの処理を加えて、別のサンプルとして用いる方法である。この方法により、サンプルの数を擬似的に10倍から100倍程度に増やすことが可能である。しかし、この手法も万能ではなく、擬似的にサンプルを増やすだけであり、サンプルに含まれる欠陥の種類などが増えるわけではないことに注意が必要である。

## 特徴量の設計

特徴量の役割は、なるべくうまく対象物を識別するための特徴を取り出すことにあり、画像認識の分野で非常に多数の方法が提案されている。すべてに万能な特徴量はなく、認識したい対象や、準備できる学習データの数などの条件に応じて適切なものを選択していく必要がある。

最も簡単には、Sobelフィルタなどでエッジ検出後の画像の輝度を特徴量とする方法や、HOG (Histogram of Oriented Gradients) 特徴、SIFT (Scaled Invariance Feature Transform) 特徴、SURF (Speeded-Up Robust Features) 特徴など、エッジ方向のヒストグラムを用いる特徴量がよく使われる。また、機械学習の一種である多層のニューラルネットを用いる、Convolutional Neural Networkでは、この特徴量設計まで含めて、学習を行うことができ、認識精度が高いという特徴がある。ただし、その分大量の学習データを要することに注意が必要である。

## 機械学習アルゴリズム

機械学習のアルゴリズムとしては、多数の方法が提案されている。最も単純なものから、ロジスティック回帰、サポートベクターマシン、ランダムフォレスト、Gradient Boosting Decision Tree、ディープラーニングと呼ばれている多層のニューラルネットなどがある。こちらも特徴量と同様に、万能なものは存在しないため、課題や条件に応じて適切なものを選択する必要がある。選択の指針としては、それぞれの方法の、学習に必要なデータ数、学習・識別速度、学習したデータ以外への汎化性能などがある。一般的に、表現能力の高い機械学習のモデルほど、大量のデータがない場合には、学習データへ過学習し、性能が出ないことも多い。

欠陥検査においては、一般的に製造される物は欠陥品と比べて良品が多いため、欠陥とされるサンプルを大量に準備するのは難しく、かつ欠陥の