

# 音から物を判別する 人工知能の作り方

牧野 浩二, 西崎 博光, 澤田 直輝

袋を振ったときの音からお菓子を判定する人工知能をうまく育てていきます。

## ● 使用したマイク

音を録音するときにはマイクの性能が影響するかどうかを調べてみましょう。実験として写真1のデバイスで音を録音します。これを写真1のように配置しました。

## ● 音の長さ

お菓子は6種類用意し、写真2のように横方向に振りました。学習用には5分、テスト用には1分の長さのデータを用意しました。実際には5種類のマイクのスタート・ボタンを押す時間も考慮して、最初の30秒は使わないデータとしました。そこで学習用のデータは5分30秒、テスト用のデータは1分30秒録音しました。

## 正答率の高いマイクの選び方

5種類のマイクで録音した結果の正答率を表1にまとめました。表からは、学習データを作ったときと同じマイクを使ってテスト・データを作ると、正答率が上がっていることが分かります。

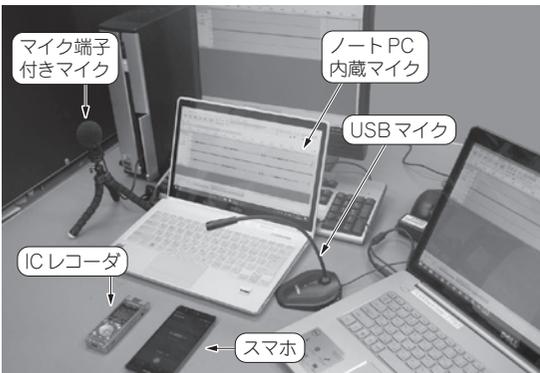


写真1 録音時のマイクの配置

例えば、PCの内蔵マイクで録音したデータを使って学習した場合は、PCの内蔵マイクで録音したテスト・データを分類した場合で、87.8%の認識率になります。しかし、ICレコーダを使って録音したテスト・データの場合には、30.6%になってしまいます。

つまり、「学習データ生成時と同じマイクで判定(推論)」することが重要です。また、どのマイクを使っても、正答率はさほど変わらないことも分かりました。

## 正答率の高い音の出し方

### ● 横に振っていたものを縦に振ってみる

お菓子の振り方を変えてみたらどのようになるのか実験をしてみました。先ほどまでは横に振っていましたが、比較として写真2のように、縦にお菓子を振りました。テスト・データは先ほどと同じように1分30秒録音し、最初の30秒は使わないものとしました。

### ● 結果

正答率を表2にまとめました。なお、数字の横に書いた▼は認識率が下がったことを表し、付いていない部分は認識率が上がったことを示しています。全部の比較をしましたが、重要な点は、学習用データを作ったときのマイクで判定を行った際の正答率の低下です。全てのマイクで、正答率が下がっています。

つまり、「サンプルは学習時と同じように振るのがよい」、裏を返せば「学習用データを作るときはできるだけ実物、実行動から生成した方がよい」ということです。

表1 マイクの種類は正答率 [%] に影響する

判定用データ 録音マイク	マイク 端子	USB マイク	内蔵 マイク	IC レコーダ	スマホ
学習用データ 録音マイク					
マイク端子	85.7	21.2	50.6	21.7	28.5
USBマイク	35.4	86.3	18.4	19.3	29
内蔵マイク	24.4	16.2	87.8	30.6	35.6
ICレコーダ	35.4	17.3	50.2	87.4	25.7
スマホ	37.6	16.9	29.3	39.2	79.3