

第3回バイオ関連スキル標準策定委員会  
2019年2月25日

## データサイエンス人材育成講座（アンケート評価）

東京医科歯科大学 医療データ科学推進室

神沼英里、田中博

時期	内容	スライドページ番号
～12月末	e-Learning教材の提供環境テスト (Google Colaboratory)	
～2月11日	教材公開ウェブサイト	2
2月12～19日	実習教材＆評価アンケート収集	3
2月25日	評価アンケートの結果	4

# ①教材公開ウェブサイトを作成

データサイエンティスト  
人材育成公開講座19-02

このサイトを検索

講座の概要  
実習教材  
サイトマップ

実習教材

<実習教材の使い方>

1. Chromeブラウザ(推奨)で、Googleアカウントへログインします。  
※Chromeアプリをインストールすればスマホ(Android)でも実習可能
2. Google Colabの初心者の方は、A1801の映像(MDDSC\_A1801.mp4)を視聴します。
3. ウェブブラウザ上でCtrl+N押下して、同じブラウザをもう一つ開きます。  
※同一ブラウザで、新規タブを開くCtrl+Tもあり。
4. M1805の映像(MDDSC\_M1805.mp4)を一つのブラウザで視聴しながら、  
もう一つのブラウザで、MDDSC\_M1805.ipynbリンクからGoogle Colabを開きます。
5. ipynbプログラムの実行を最初から最後まで進めます。
6. 本教材について、質問を受付致します。Googleアカウントでログインの状態で、  
このページ下にコメント欄がありますので、直接書き込んで下さい(時々確認して、回答致します)。
7. 終了後に、教材評価用アンケートの記入をお願い致します。

<教材概要>

\* タイトル: 深層学習による画像分類と前処理効果

\* 概要: オープンデータのマウスOCT画像を用いて、疾患と正常を分類する深層学習モデルを構築します。更に前処理の種類別の分類精度の変化を、Google ColabのPythonコードを違いながら解説します。

<実習環境>

\* Google Colab = クラウド実行環境  
\* Python = プログラミング言語  
\* Keras, Tensorflow = 深層学習ライブラリ

## <プログラミング実習用教材>

番号	タイトル	教材概要	教材LINK
M1804	実習教材の使い方	* 実習の使い方 * コメントの作成 * アンケートの記入	<映像> * <a href="#">MDDSC_M1804.mp4</a> 03m08s
M1805	Google Colabの基本操作	* コードの保存 (ファイルコピー、名前変更) * Colab環境の確認	<プログラムコード> * <a href="#">MDDSC_M1805.ipynb</a>  <映像> * <a href="#">MDDSC_M1805.mp4</a> 11m06s
M1806	深層学習の概要	* 深層学習とは * CNNモデル	<映像> * <a href="#">MDDSC_M1806.mp4</a> 16m09s
M1807	深層学習の画像分類 [1] GPU/CPUの計算時間	* GPU/CPUの切替 * 深層学習 * MNIST画像データセット	<プログラムコード> * <a href="#">MDDSC_M1807.ipynb</a>  <映像> * <a href="#">MDDSC_M1807.mp4</a> 12m17s
M1808	深層学習の画像分類 [2] 前処理と分類精度	* 光干渉断層撮影 (OCT) 画像データ * 深層学習 * データ拡張 (前処理)	バグ修正版を公開しました。本ページ下とipynb中に説明を記入しました。2019.2.15  <プログラムコード> * <a href="#">MDDSC_M1808修正版.ipynb</a>  <映像> * <a href="#">MDDSC_M1808修正版sub1.mp4</a> 10m55s * <a href="#">MDDSC_M1808修正版sub2.mp4</a> 05m21s

Google Drive

MDDSC\_M1808修正版.ipynb

https://colab.research.google.com/github/190202/colab/blob/master/MDDSC\_M1808修正版.ipynb

MDDSC\_M1808修正版.ipynb

ファイル 編集 表示 挿入 ランタイム ツール ヘルプ

コピート プリント 終了

ELHNet: マウス蝸牛の内リンパ水腫検出  
(文献番号=PMID:2868782)

ELHNet: a convolutional neural network for classifying coc endolymphatic hydrops imaged with optical coherence tomography

光干渉断層撮影 (Optical Coherence Tomography: OCT)

(a) (b) Normal Hydrops

E

種類 : 動画

サイズ 45 MB (46,942,500 バイト)

MP4動画

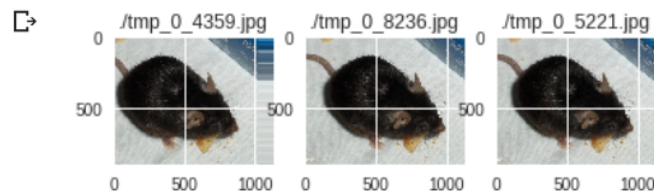
## Google Colab初心者対象

番号	タイトル	教材概要	教材LINK
A1801	Google Colaboratory の開始方法	* Google Colabとは * GPU、Cloud環境	<映像> * <a href="#">MDDSC_A1801.mp4</a> 04m31s

## ②実習教材と評価アンケート収集

### \* 実習教材：深層学習による画像分類と前処理効果

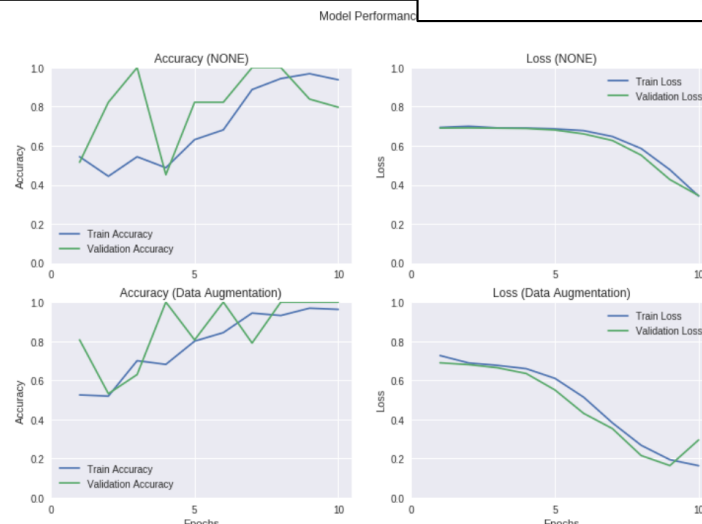
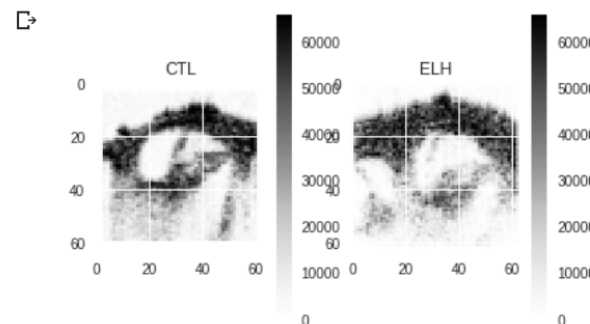
```
[ ] 1 #-----[2-4:データ拡張の画像を視覚化]-----  
2 #  
3 # データ拡張で生成した、3つの画像（ランダムな平行移動画像）を表示する。  
4 #  
5 #-----  
6  
7 from glob import glob  
8 import matplotlib.pyplot as plt  
9  
10 files = glob('./*.jpg');  
11 i=1  
12 for filename in files:  
13     out=load_img(filename)  
14     plt.subplot(1,3,i)  
15     plt.title(filename)  
16     plt.imshow(out)  
17     i += 1  
18     if i > 3:  
19         break  
20  
21 plt.show()
```



Layer (type)	Output Shape	Param #
conv2d_7 (Conv2D)	(None, 62, 62, 64)	640
activation_19 (Activation)	(None, 62, 62, 64)	0
max_pooling2d_7 (MaxPooling2)	(None, 31, 31, 64)	0
flatten_7 (Flatten)	(None, 61504)	0
dense_13 (Dense)	(None, 1024)	62981120
activation_20 (Activation)	(None, 1024)	0
dropout_7 (Dropout)	(None, 1024)	0
dense_14 (Dense)	(None, 2)	2050
activation_21 (Activation)	(None, 2)	0
Total params: 62,983,810		
Trainable params: 62,983,810		
Non-trainable params: 0		

```
[ ] 14 import numpy as np  
15  
16 #-----サンプル画像読み込み-----  
17  
18 im_ctl = Image.open("../ELH64x64small/Test/CTL/1_160.png") #CTL画像の読み込み  
19 im_elh = Image.open("../ELH64x64small/Test/ELH/2_160.png") #ELH画像の読み込み  
20  
21 #-----サンプル画像表示-----  
22 plt.subplot(1,2,1)  
23 plt.title('CTL')  
24 plt.imshow(np.array(im_ctl))  
25 plt.axis('image')  
26 plt.colorbar()  
27  
28 plt.subplot(1,2,2)  
29 plt.title('ELH')  
30 plt.imshow(np.array(im_elh))  
31 plt.axis('image')  
32 plt.colorbar()  
33 plt.show()
```

\* OCT images of mice cochleae.  
\* CTL vs ELH(蝸牛内リンパ水腫)



### 教材のフィードバック

次年度カリキュラム組立ての参考情報をお知らせ頂ければ幸いです。(2019年2月8日/木)

\*必須

あなたの背景情報をお知らせください\*

☐ 博士人材 (学生、ポスドク、etc)

☐ 企業人材

教材内容について\*

コメントや改善点の記載をお願い致します。

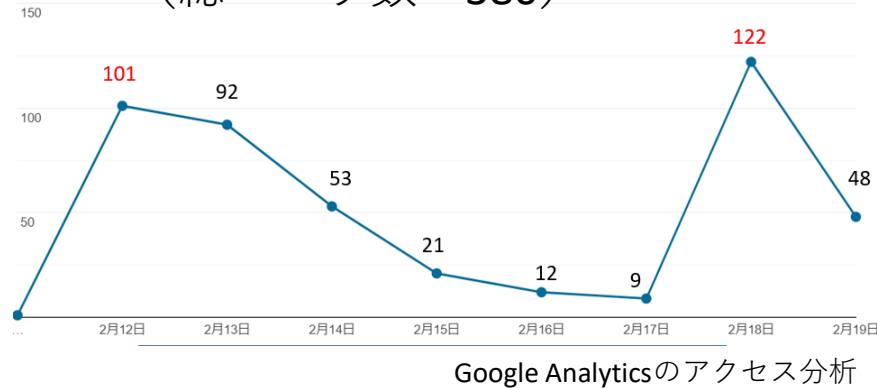
回答を入力

\* 評価アンケート (Google Form)

送信

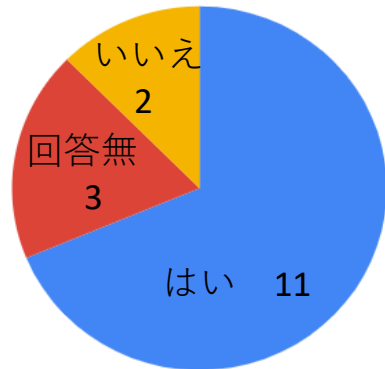
## ②評価アンケートの結果

### ■教材サイトのアクセスユーザ数 (総ユーザ数：380)



### ■次年度参加の意向(11/16=約7割)

5. 来年も同様の講座が開設されれば参加してみたいとお考えですか。もしくは同僚に進めてみようとお考えですか？



#### いいえ 2 件のコメント

現状では、google検索で調べれば無料でわかる範囲だと感じた。パソコンの動作原理やプログラミングが全くわからない人(検索しても内容を理解できない人)が、オープンソースのtutorialを応用して現場の解析に活用できるまで育つ教育システムであれば、導入しようと考えられる。

どのような人に、どんなスキルを新たに身に付けてもらうための導入講座なのかが中途半端な印象。短時間のe-learningという制約がある以上、例えば深層学習の素養がある方に画像解析を解説するなら、基礎理論やMNISTの例は不要な気がしますし、逆に初学者対象なら、SVMやランダムフォレスト等の機械学習ライブラリーの利用程度で、プログラム文の意味するところが概ね理解できるレベルで進めた方が意味のある講座になるのではと感じました。

### ■ユーザの背景情報

2. あなたの背景情報をお知らせください	人数
アカデミア (学生、ポスドク)	3
企業 (管理職)	1
企業 (非管理職、研究員など)	12
総計	16

### ■コメント

深層学習自体が初めてだったので、導入のスライドがあり分かり易かった。

DeepLearningの画像解析の流れをみることで、プログラムコードをみながら説明を聞くことができたこと、資料と説明が分かり易かったです。今後の勉強のきっかけになりました。ありがとうございます。

データ拡張による効果を知ることができて良かったです。

大抵の教科書に載っているMNISTだけでなく、ライフサイエンス研究に関わる画像解析の例を用いて深層学習の体験ができたことは、学習を進めるモチベーションになった。Google Colabを用いてクラウド上でJupyter Notebookを簡単に操作できることが判り、さらにプログラムの動作を実感できたことも、学習の入り口として良い経験となった。

深層学習での過学習への対応、データ拡張についての概要を知ることができて、ハンズオンで一通り実習でき勉強になった。

講義資料が丁寧に作成されていたので、自分のペースで読み返すことができました。

実際の操作方法を学べ良かったです

良かった点

実務に応用するには、改めて相当の学習が必要であることは自明。相当に作りこんだ内容なので、モチベーションだけで終わっては勿体無い。やる気が出た人向けに、今後の学習教材(参考書、学習サイトなど)の紹介があっても良かった気がした。

構築したモデルの評価法、精度向上のために大きく構造を変えないといけないのか、データ数を増やすのがいいのか、などそれぞれの判断を下すときの着眼点が知れるととってもよかった。

ノートパソコン上で2つのウィンドウを用いて閲覧すると、動画の文字が小さくて見にくくなってしまいました。

導入教材として不満はありません

改善点