Final Project

Age prediction from handwritten document images

Due date: 4/07/2022, 23:59

Frontal check: 6/07/2022, 10:00-13:00

The objective of the course's final project is to train a neural network to predict age from handwriting samples.

Age prediction from the handwritten sample is of great interest in several areas, including psychology, historical document analysis, and handwriting biometrics. Previous studies on the correlation between age and handwriting found that aging alters the behavior of the features extracted from handwriting, including an increase in the number of pen lifts, a decrease in writing speed, a reduction in pen pressure, an irregular writing rhythm, irregular character shapes and slope, and a loss of smoothness in the trajectory.

The dataset

You will run the experiments on the KHATT dataset [1]. The KHATT (KFUPM Handwritten Arabic Text) dataset 40] consists of 1000 handwritten forms written in Arabic by 1000 writers. The labels contain gender (male/female) and age. The age is assigned into one of four groups:

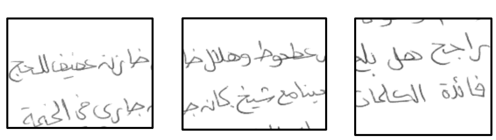
1. group_1: "< 15"
2. group_2: "16 – 25"
3. group_3: "26 −50"
4. group_4: "> 50".

The forms were filled in mostly by high school and university students. Each writer contributed four paragraphs.

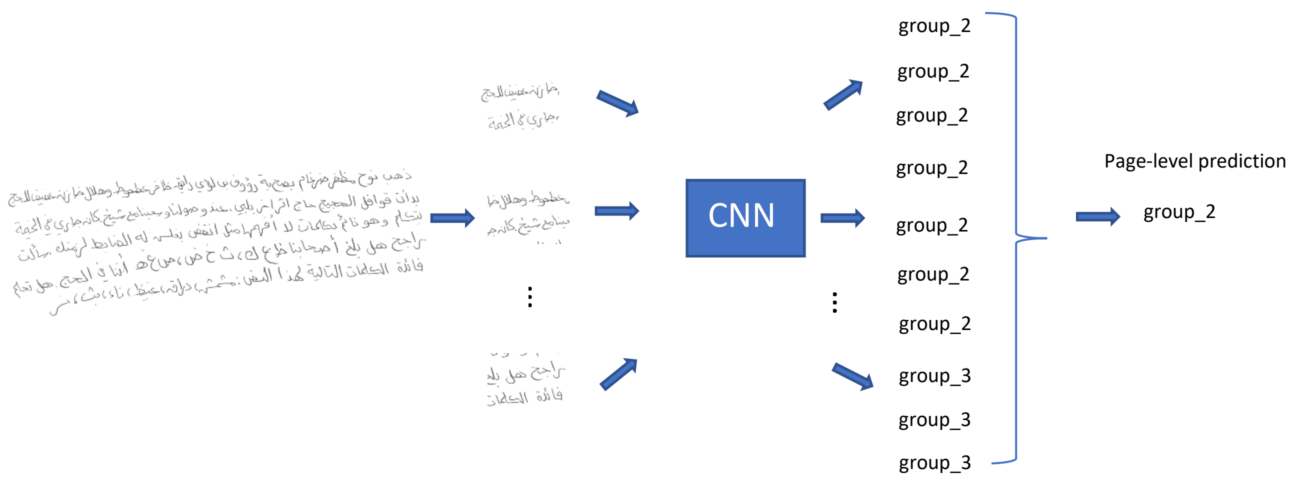
In this project, you will work with the subset of KHATT that consists of samples of 270 writers belonging to groups 2 and 3 ("16-25" and "26-50"), with 135 writers in each group and four paragraphs from each writer. You will be provided with the train/test split for each group. There is no validation set.

Experimental settings

Training. In each experiment, the model will be trained on patches extracted from document images, i.e., an input to the model is an image patch. The patches should be extracted by moving a sliding window of size 400 × 400 pixels at a stride of 200 pixels in vertical and horizontal directions. The number of patches extracted from each page depends on the text written on the page and can vary from page to page. The following figure presents the examples of the extracted patches.



Testing. For prediction on *test* set, the manuscript image should be also cut into overlapping 400×400 pixels patches at a stride of 200 pixels. Classify each patch and then assign the resulting *page-level label* by the majority voting scheme over all patches from the same page. For example, if there are 10 patches extracted from the page, and for 7 patches the classifier returned the prediction "age group_2" and for the rest 3 patches the prediction is "age group_3", then the final prediction for this page is "age group_2".



Experiment 1

In the first experiment your task is to build the following neural network (like the network in homework 3):

INPUT=>[CONV=>RELU=>CONV=>RELU=>POOL=>DO]*3=>FC=>RELU=>DO=>FC

CNN layers:

Use filter of size 3×3, padding = 'same'

First iteration:

CONV=>RELU=>CONV=>RELU=>POOL=>DO

32 filters in each CONV

DO with p = 0.25

Second iteration:

CONV=>RELU=>CONV=>RELU=>POOL=>DO

64 filters in each CONV

DO with p = 0.25

Third iteration:

CONV=>RELU=>CONV=>RELU=>POOL=>DO

128 filters in each CONV

DO with p = 0.25

Pooling layers:

Use max pooling of size 2×2

Layer FC=>RELU=>DO

FC – 512 neurons

DO with p = 0.5

The last FC layer (output)

One neuron – this is the binary problem; we are trying to predict only two classes. Add the sigmoid activation function to the output layer.

! Use *binary_crossentropy* with `model.compile`, which is purpose-built for binary classifiers.

Experiment 2

In this experiment, you will implement transfer learning (fine-tuning) with ImageNet pre-training.

You should choose two networks from the following list – any that you like. All these networks are implemented in keras.

- VGG16
- ResNet50
- Xception
- EfficientNet

After you have chosen the models, train them:

a) Load the model without the top layer
b) Freeze the layers
c) Add a classifier on top of the convolutional base - add a fully connected layer followed by a sigmoid with one outputs
d) Train the model for 20 epochs
e) Unfreeze the earlier layers and train the model for additional 20 epochs
f) Save the model

After training, report the results on *test* set for each model:

1. Overall average accuracy
2. Average accuracy for each class separately
3. Confusion matrix

Results

Which accuracy should you achieve? Honestly, I don't know 😊. To the best of my knowledge, nobody applied neural network on the KHATT dataset for age prediction. All the studies I found in the literature, applied traditional machine learning techniques. The following table lists the results of other studies.

Method	Accuracy (%)	Study	Settings
Hu moments + k-means clustering	64.44	[2]	Two age groups: 2 and 3
SVM	55.55	[3]	Three age groups: 1, 2 and 3
SVM	67.78	[4]	Two age groups: 2 and 3
SVM, MIN-MAX	81.11	[5]	Two age groups: 2 and 3

In this project, you will check whether CNN models outperform traditional methods for age classification.

Submission instructions

You should submit a zip file with the following files inside:

1. Your code with all the experiments (all the stages should be included in the code)
2. Trained models from each experiment in 'h5' files (three models in total: one from experiment 1, and two from experiment 2)
3. "Results.txt" file in which results of all experiments are summarized (for each model report the overall accuracy, the accuracy for each group, and the confusion matrix)
4. "readme.txt" file

References

1. Mahmoud, S.A., Ahmad, I., Al-Khatib, W.G., Alshayeb, M., Parvez, M.T., M ̈argner, V., Fink, G.A.: KHATT: An open Arabic offline handwritten text database. Pattern Recognition 47(3), 1096–1112 (2014). https://doi.org/10.1016/j.patcog.2013.08.009
2. Basavaraja, V., Shivakumara, P., Guru, D.S., Pal, U., Lu, T., Blumenstein, M.: Age estimation using disconnectedness features in handwriting. In: 2019 International Conference on Document Analysis and Recognition (ICDAR). pp. 1131–1136 (2019). https://doi.org/10.1109/ICDAR.2019.00183
3. Bouadjenek, N., Nemmour, H., Chibani, Y.: Age, gender and handedness prediction from handwriting using gradient features. In: 2015 13th International Conference on Document Analysis and Recognition (ICDAR). pp. 1116–1120 (2015). https://doi.org/10.1109/ICDAR.2015.7333934

4. Bouadjenek, N., Nemmour, H., Chibani, Y.: Histogram of oriented gradients for writer's gender, handedness and age prediction. In: 2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA). pp. 1–5. IEEE (2015). https://doi.org/10.1109/INISTA.2015.7276752
5. Bouadjenek, N., Nemmour, H., Chibani, Y.: Robust soft-biometrics prediction from off-line handwriting analysis. Applied Soft Computing 46, 980–990 (2016). https://doi.org/10.1016/j.asoc.2015.10.021