

Gender classification problem:

By Dina Maizlis

Abstract:

By analyzing the face we get a lot of information such as age, gender, ethnicity, identity, feeling and more. A gender classification system uses face of a person from a given image to tell the gender (male/female) of the given person. A successful gender classification approach can boost the performance of many other applications including face recognition and smart human-computer interface. This article illustrates the general processing steps for gender classification based on frontal face images. This article contains three different models for gender classification problem - Logistic Regression, Neural Network (NN) and Convolutional Neural Network (CNN)

Introduction:

In recent years with the development of technology the identified by an image becomes part of the everyday. Deepening the problem of gender classification contributes to the field of computer vision and the field of data analysis.

Explain- In the field of computer vision: improve existing systems that work on identifying a person by image for example AppellID- do actions such as opening a phone only if it identifies the user and check at the entrance to the country that the passport belongs to the holder.

In the field of data analysis- through gender identification it is possible to adjust the information a person receives to be relevant to him. For example matching advertisements while using social networks.

This article will try to answer the question of gender classification by a face image, because this is the key characteristic feature in human.

Related work:

This issue of gender classification has been researched in the past (since 1991) and there are many articles.

- Article Gender Classification Using Proposed CNN-Based Model and Ant Colony Optimization <https://www.mdpi.com/2227-7390/9/19/2499>
- Kaggle notebook- <https://www.kaggle.com/drfrank/face-image-classification/notebook>

Required background:

Logistic regression- is a classification algorithm used to find the probability of event success and event failure. It is used when the dependent variable is binary(0/1, True/False, Yes/No) in nature. It supports categorizing data into discrete classes by studying the relationship from a given set of labelled data. It learns a linear relationship from the given dataset and then introduces a non-linearity in the form of the Sigmoid function.¹

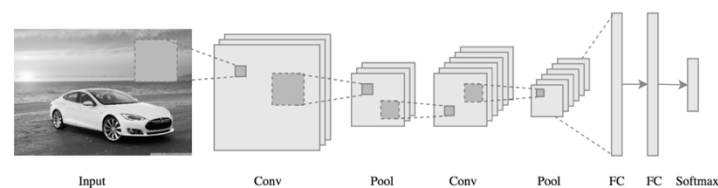
¹ <https://www.geeksforgeeks.org/advantages-and-disadvantages-of-logistic-regression/>

Neural Network (NN)- A neural network (NN), in the case of artificial neurons called artificial neural network (ANN) or simulated neural network (SNN), is an interconnected group of natural or artificial neurons that uses a mathematical or computational model for information processing based on a connectionist approach to computation.²

Convolutional Neural Network (CNN)- In deep learning, a convolutional neural network (CNN) is a class of artificial neural network, most commonly applied to analyze visual imagery. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on the shared-weight architecture of the convolution kernels or filters that slide along input features and provide translation equivariant responses known as feature maps

A convolutional neural network consists of an input layer, hidden layers and an output layer. In any feed-forward neural network, any middle layers are called hidden because their inputs and outputs are masked by the activation function and final convolution. In a convolutional neural network, the hidden layers include layers that perform convolutions.³

4



Convolutional layers- Convolutional layers are the layers where filters are applied to the original image, or to other feature maps in a deep CNN. This is where most of the user-specified parameters are in the network. The most important parameters are the number of kernels and the size of the kernels.⁵

Max-Pooling layers- Max pooling is a pooling operation that selects the maximum element from the region of the feature map covered by the filter. Thus, the output after max-pooling layer would be a feature map containing the most prominent features of the previous feature map.⁶

Dropout layer- Dropout is a technique where randomly selected neurons are ignored during training. They are “dropped-out” randomly. This means that their contribution to the activation of downstream neurons is

² https://en.wikipedia.org/wiki/Neural_network

³ https://en.wikipedia.org/wiki/Convolutional_neural_network#Definition

⁴ <https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/>

⁵ <https://towardsdatascience.com/simple-introduction-to-convolutional-neural-networks-cdf8d3077bac>

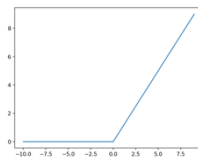
⁶ <https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/>

temporally removed on the forward pass and any weight updates are not applied to the neuron on the backward pass.^{7 8}

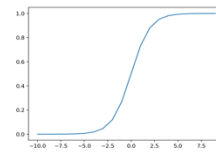
activation layer -An activation function in a neural network defines how the weighted sum of the input is transformed into an output from a node or nodes in a layer of the network.

ReLU activation layer- This is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. The rectified linear activation function overcomes the vanishing gradient problem, allowing models to learn faster and perform better.⁹

Sigmoid activation layer- The function takes any real value as input and outputs values in the range 0 to 1. The larger the input (more positive), the closer the output value will be to 1.0, whereas the smaller the input (more negative), the closer the output will be to 0.



ReLU function- $\max(0.0, x)$



Sigmoid function $-1.0 / (1.0 + e^{-x})$

project description:

Library- numpy, pandas, matplotlib, seaborn, plotly, sklearn, tensorflow and keras.

Dataset- contains 23075 different images, of size 48*48 and labels.

(link for the dataset- <https://www.kaggle.com/nipunarora8/age-gender-and-ethnicity-face-data-csv>)

Train/Test/Validation- the data divided to train 0.7% test 0.2% and validation 0.1%.

CNN model - Used four types of layers:

- *convolution layer*- A convolution layer transforms the input image in order to extract features from it.
- *Max pooling layer*- is a pooling operation that calculates the maximum value for patches of a feature map, and uses it to create a down sampled feature map. It is usually used after a convolutional layer.
- *dropout layer* - is a regularization technique for reducing overfitting in artificial neural networks by preventing complex co-adaptations on

⁷ <https://machinelearningmastery.com/dropout-regularization-deep-learning-models-keras/>

⁸ The image taken from a presentation learned in a lecture

⁹ <https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/>

training data. It is an efficient way of performing model averaging with neural networks. The term dilution refers to the thinning of the weights

- **dense layer**- A Dense layer feeds all outputs from the previous layer to all its neurons, each neuron providing one output to the next layer. It's the most basic layer in neural networks

Model: "sequential"		
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 46, 46, 32)	320
max_pooling2d (MaxPooling2D)	(None, 23, 23, 32)	0
conv2d_1 (Conv2D)	(None, 21, 21, 32)	9248
max_pooling2d_1 (MaxPooling2D)	(None, 10, 10, 32)	0
dropout (Dropout)	(None, 10, 10, 32)	0
conv2d_2 (Conv2D)	(None, 8, 8, 64)	18496
max_pooling2d_2 (MaxPooling2D)	(None, 4, 4, 64)	0
flatten (Flatten)	(None, 1024)	0
dense (Dense)	(None, 64)	65600
dropout_1 (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 1)	65
Total params: 93,729		
Trainable params: 93,729		
Non-trainable params: 0		

This model contains the layers: convolution layer- size 3*3, max pooling- size 2*2, convolution layers- size 3*3, max pooling- size 2*2, dropout layer- size 0.3%, : convolution layer- size 3*3, max pooling- size 2*2, flatten , dense, dropout layer- size 0.5% and dense.

All the layers used activation function of "ReLU" except the last layer that used activation function of " Sigmoid".

The accuracy of this models is 0.899.

Previous attempts:

The previous attempts include experience of different types of layers in different combination and different activation function.

And include experience of number of epochs, batch Size and validation.

Experiments/simulation results:

The experiments results of Convolutional Neural Network (CNN) is accuracy of 0.899. During the simulation , with different layers and different activation function the accuracy was in range of 0.72-0.864.

The experiments results of Logistic regression is accuracy of 0.83, and Neural Network (NN) is accuracy of 0.849.

Conclusions:

All the models presented in the article are better models than the random.

The CNN models is the model that got the best accuracy of 0.899.

- for the codes and more information: <https://github.com/dinamaizlis/gender-classification>