# Helwan University Faculty of Computers & Artificial Intelligence – General Programme Computer Science Department – Module: CS361 Artificial Intelligence – Fall “Semester 1” 2021-2022

# 15 ) An Automated Optical Character Recognition of Handwritten English Letters using Artificial Neural Networks,. Optical character recognition or optical character reader (OCR) is the automated conversion of images of typed, handwritten, or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (e.g., the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (e.g., from a television broadcast). Build an OCR using Artificial Neural Networks to identify Handwritten English Letters. You MAY use the following data-set: https://www.kaggle.com/crawford/emnis

# Team members

1-عمر حامد عبد النبي سند

Id -201900505 team leader

Email: [omarhamed200160@gmail.com](mailto:omarhamed200160@gmail.com) phone : 01060770592

2-محمود سعد محمد بكر

Id-201900768

3-عبدالرحمن ابراهيم عبدالرحمن محمد

Id-201900397

4-صلاح بكر فاروق عبدالله

Id-201900382

5-اسماء محمود عبدالجبار محمود

Id-201900150

6- مدحت طه محمداحمد

Id-20180575

# EMNIST (Extended MNIST)

# An extended variant of the full NIST dataset

License

[CC0: Public Domain](https://creativecommons.org/publicdomain/zero/1.0/)

# **(1) Project idea in details**

# Description

EMNIST

The MNIST dataset has become a standard benchmark for learning, classification and computer vision systems. Contributing to its widespread adoption are the understandable and intuitive nature of the task, its relatively small size and storage requirements and the accessibility and ease-of-use of the database itself. The MNIST database was derived from a larger dataset known as the NIST Special Database 19 which contains digits, uppercase and lowercase handwritten letters. This paper introduces a variant of the full NIST dataset, which we have called Extended MNIST (EMNIST), which follows the same conversion paradigm used to create the MNIST dataset. The result is a set of datasets that constitute a more challenging classification tasks involving letters and digits, and that shares the same image structure and parameters as the original MNIST task, allowing for direct compatibility with all existing classifiers and systems. Benchmark results are presented along with a validation of the conversion process through the comparison of the classification results on converted NIST digits and the MNIST digits..

Format

There are six different splits provided in this dataset and each are provided in two formats:

1. Binary (see emnist*source*files.zip)
2. CSV (combined labels and images)
   * Each row is a separate image
   * 785 columns
   * First column = class\_label (see mappings.txt for class label definitions)
   * Each column after represents one pixel value (784 total for a 28 x 28 image)

ByClass and ByMerge datsets

The full complement of the NIST Special Database 19 is available in the ByClass and ByMerge splits. These two datasets have the same image information but differ in the number of images in each class. Both datasets have an uneven number of images per class and there are more digits than letters. The number of letters roughly equate to the frequency of use in the English language.

* train: 697,932
* test: 116,323
* total: 814,255
* classes: ByClass 62 (unbalanced) / ByMerge 47 (unbalanced)

Balanced dataset

The EMNIST Balanced dataset is meant to address the balance issues in the ByClass and ByMerge datasets. It is derived from the ByMerge dataset to reduce mis-classification errors due to capital and lower case letters and also has an equal number of samples per class. This dataset is meant to be the most applicable.

* train: 112,800
* test: 18,800
* total: 131,600
* classes: 47 (balanced)

Letters datasets

The EMNIST Letters dataset merges a balanced set of the uppercase and lowercase letters into a single 26-class task.

* train: 88,800
* test: 14,800
* total: 103,600
* classes: 37 (balanced)

Digits and MNIST datsets

The EMNIST Digits and EMNIST MNIST dataset provide balanced handwritten digit datasets directly compatible with the original MNIST dataset.

* train: Digits 240,000 / MNIST 60,000
* test: Digits 40,000 / MNIST 10,000
* total: Digits 280,000 / MNIST 70,000
* classes: 47 (balanced)

we can use it to

1-Getting to know the contagious oases

2-Getting traffic fines

3-Translation

4-used in windows 10

5-Read barcodes and use them in all fields

(2) Main functionalities

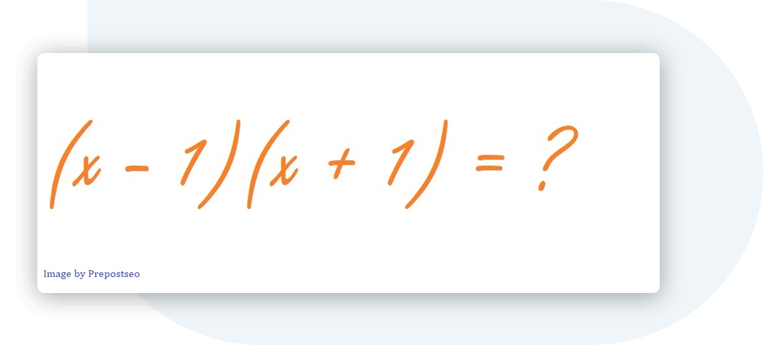
Let’s explore some of the classic features of this image to text app.

### **1. Extract text from low-resolution images**

Students often take photos of book pages and notes for exams. These pictures are usually blurred due to poor camera quality.

The text extractor can take out text from low-resolution and blurry images as well.

### **2. Identify math equations**



You may have some pictures of algebraic or geometric formulas if you are a math geek. The good news is, this OCR tool not only gets simple text for you but also extracts complex mathematical equations like a pro.

### **3. Free of cost**

You can upload as many scanned books, notes, images, and photos as you want.

Our free software extract correct text each time without asking you to sign up or paid subscription.

### **4. Trustworthy and secure**

You are safe and secure with us. We always aim and promise to keep your data safe and secure. We adhere to our policy that we will neither pass your data to any third party nor it will be stored in our database.

### **5. Multiple language support**

The image-to-text ocr provides multilingual support. It can translate images into text in more than 30 languages.

All supported languages are mentioned at the end.

### **6. Extract text via URL**

Apart from uploading, a picture can be converted to text by inserting the URL of the pic into the URL box.

This feature is very useful when you browse the Internet and come across an image that contains the text.

### **7. Download text file**

Data can be lost or misplaced if not saved appropriately. You can directly download converted text as a file instead of copying the text

(3)Similar applications in the market

there are five Android Apps to Convert Image to Text – [OCR Apps]

1.ماسح الوثائق الضوئ Camscanner

2.Google Goggles

3. OneNot

4.Microsoft Lens - PDF Scanner

5.Text Fairy (OCR Text Scanner)

(4) An initial literature review of Academic

We use OCR sum information of the ocr

what is OCR in simple words?

OCR. Stands for "Optical Character Recognition." Some OCR programs allow you to scan a document and convert it to a word processing document in a single step. While OCR technology was originally designed to recognize printed text, it can be used to recognize and verify handwritten text as well.

Why is OCR needed?

OCR can make your life easier by: Reduce or eliminate costly data entry by automatically grabbing information you need from paper and putting it where it needs to go. Enabling entirely new ways to process documents that can eliminate “human touches”, thereby reducing costs and dramatically reducing processing times.

We use ANN sum information os ann

What are Artificial Neural Networks?

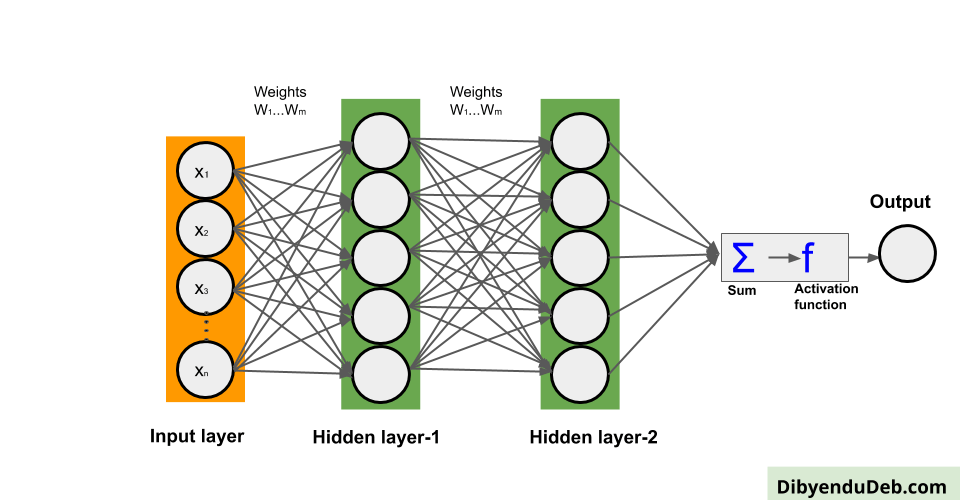
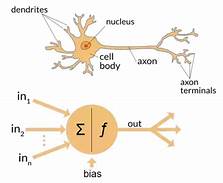
Simply put, Artificial Neural Networks are software implementations of the neural structures of human brain. ANN is a computational system influenced from the structure, processing capability and learning ability of a human brain.

Not diving deep into the complex biology of it, let us take a look at the structure of our brain. Human brain contains billions of neurons that act as organic switches. All these neurons are interconnected to form a huge and complex structure called Neural Network. The output of a single neuron is dependent on inputs from thousands of interconnected neurons.

This behavior of the brain is key to Artificial Neural Networks as they simply try to replicate this action of the brain. This can be achieved in two ways:

Supervised ANN

Unsupervised ANN



We use sum library

Numpy

What is NumPy?

NumPy is a Python library used for working with arrays.

It also has functions for working in domain of linear algebra, fourier transform, and matrices.

NumPy was created in 2005 by Travis Oliphant. It is an open source project and you can use it freely.

NumPy stands for Numerical Python.

Why Use NumPy?

In Python we have lists that serve the purpose of arrays, but they are slow to process.

NumPy aims to provide an array object that is up to 50x faster than traditional Python lists.

The array object in NumPy is called ndarray, it provides a lot of supporting functions that make working with ndarray very easy.

Arrays are very frequently used in data science, where speed and resources are very important.

Pandas

Pandas is an open-source Python Library providing high-performance data manipulation and analysis tool using its powerful data structures. The name Pandas is derived from the word Panel Data – an Econometrics from Multidimensional data.

In 2008, developer Wes McKinney started developing pandas when in need of high performance, flexible tool for analysis of data.

Pytesseract

Pytesseract or Python-tesseract is an Optical Character Recognition (OCR) tool for Python. It will read and recognize the text in images, license plates etc. Python-tesseract is actually a wrapper class or a package for Google’s Tesseract-OCR Engine.

**The references**

1. <https://ieeexplore.ieee.org/abstract/document/7966217>
2. <https://pypi.org/project/python-mnist/>
3. <https://arxiv.org/abs/1702.05373v1>.
4. <https://www.nist.gov/itl/products-and-services/emnist-dataset>
5. <https://dibyendudeb.com/artificial-neural-network/>
6. <https://www.w3schools.com/python/numpy/numpy_intro.asp>
7. <https://stackabuse.com/pytesseract-simple-python-optical-character-recognition/>
8. <https://realpython.com/pandas-python-explore-dataset/>
9. <https://www.tutorialspoint.com/python_pandas/python_pandas_introduction.htm>
10. <https://askinglot.com/what-is-ocr-in-python>
11. <https://www.techantena.com/3969/best-android-apps-convert-image-text-ocr-apps/>
12. Computer Science Department – Module: CS361 Artificial Intelligence – Fall “Semester 1” 2021-2022 prof. Amr.S.Ghoneim
13. CS395: Selected CS1 (Introduction to Machine Learning) Associate Prof. Wessam El-Behaidy

**Some of the tools used in the algorithm**

Keras Conv2D is a 2D Convolution Layer, this layer creates a convolution kernel that is wind with layers input which helps produce a tensor of outputs.

Kernel: In image processing kernel is a convolution matrix or masks which can be used for blurring, sharpening, embossing, edge detection, and more by doing a convolution between a kernel and an image.

he Python filter () method is used to filter lists, sets, and tuples

kernel\_size: An integer or tuple/list of 2 integers, specifying the height and width of the 2D convolution window. Can be a single integer to specify the same value for all spatial dimensions

The activation function is the non-linear function that we apply over the input data coming to a particular neuron and the output from the function will be sent to the neurons present in the next layer as input.

MaxPooling2D is used to max pool the value from the given size matrix and same is used for the next 2 layers. then, Flatten is used to flatten the dimensions of the image obtained after convolving it. Dense is used to make this a fully connected model and is the hidden layer. Dropout is used to avoid overfitting on the dataset.

pool\_size: a pool size is a number of the connection objects that the pool can support. If this argument is not given, the default is 5. The pool size cannot be 0 or less than 0.

Strides are the number of bytes to jump-over in the memory in order to get from one item to the next item along each direction/dimension of the array. In other words, it's the byte-separation between consecutive items for each dimension. For example: >>> a = np.arange (1,10).reshape (3,3) >>> a array ([ [1, 2, 3], [4, 5, 6], [7, 8, 9]])

n Python, for some cases, we need a one-dimensional array rather than a 2-D or multi-dimensional array. For this purpose, the numpy module provides a function called numpy.ndarray.flatten (), which returns a copy of the array in one dimensional rather than in 2-D or a multi-dimensional array. Syntax ndarray.flatten (order='C')

We can plot a density plot in many ways using python. Let’s look at a few commonly used methods. 1. Using Python scipy.stats module scipy.stats module provides us with gaussian\_kde class to find out density for a given data

* Below is the python code for it: from keras.layers.core import Dropout model = Sequential ([ Dense (output\_dim=hidden1\_num\_units, input\_dim=input\_num\_units, activation='relu'), Dropout (0.25), Dense (output\_dim=output\_num\_units, input\_dim=hidden5\_num\_units, activation='softmax'), ])

The summary () method is used to obtain a table which gives an extensive description about the regression results Syntax : statsmodels.api.OLS (y, x) Parameters : y : the variable which is dependent on x x : the independent variable Code: