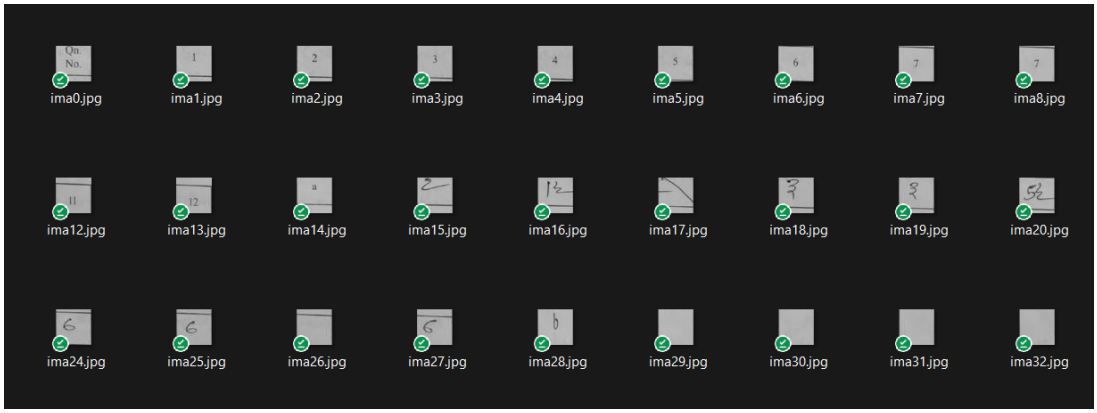
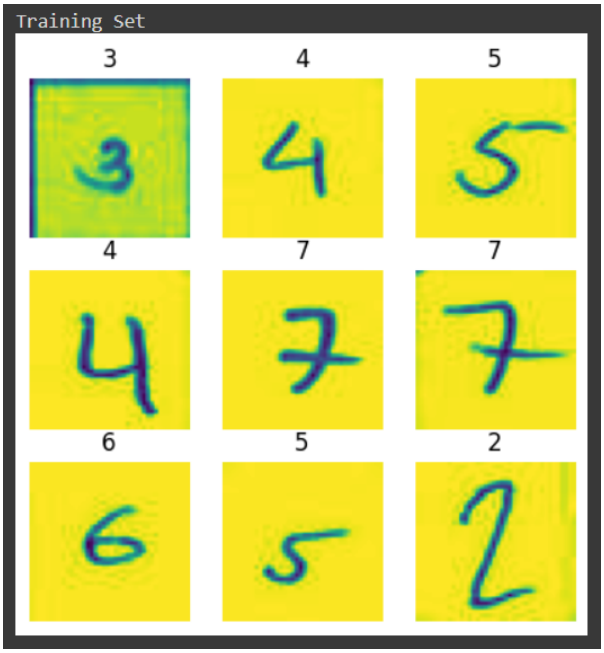
Supplementary Material for MARKS2CSV Conference Paper (ICTEST 2023)



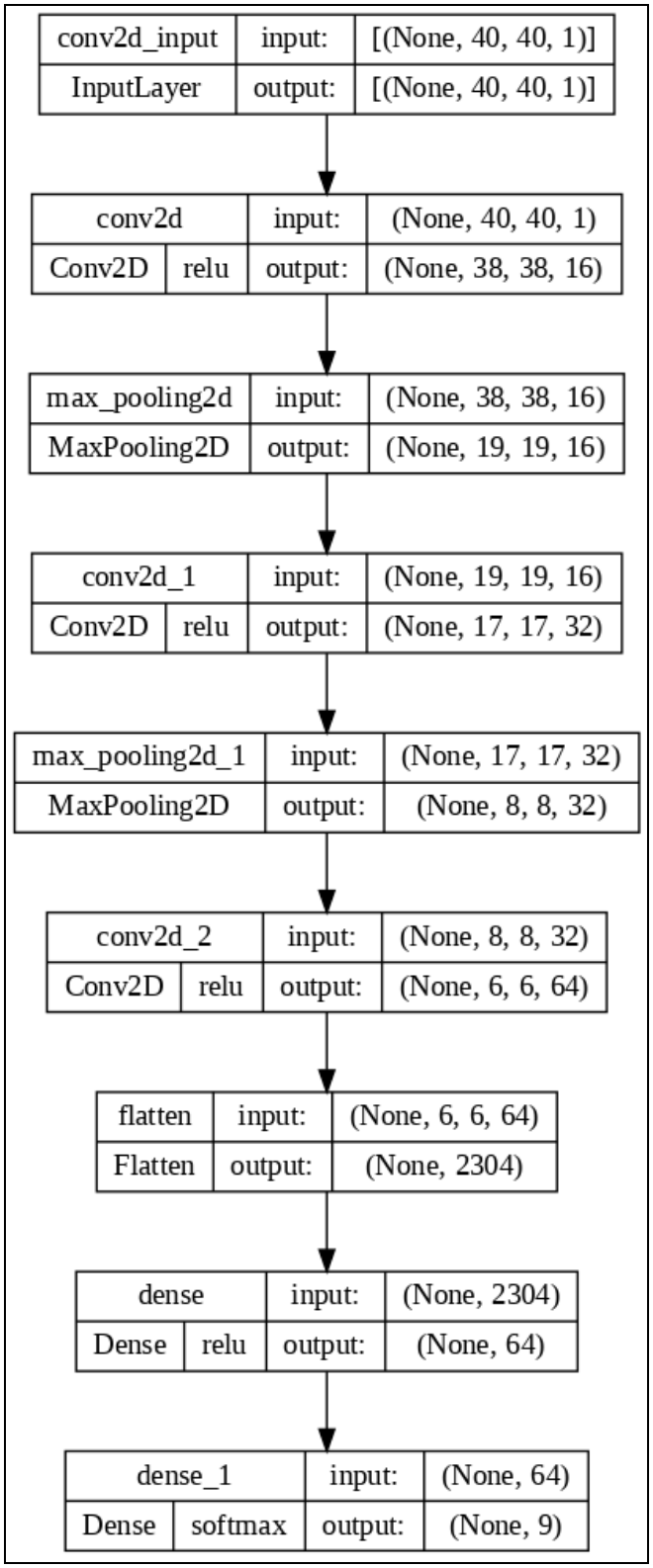
Set of extracted cells of handwritten digits from PDF



Snapshot of the training set with ground truth values

```python
# create the neural network
model = tf.keras.Sequential([
    tf.keras.layers.Conv2D(16, 3, activation='relu', input_shape=(40, 40, 1)),
    tf.keras.layers.MaxPooling2D(),
    tf.keras.layers.Conv2D(32, 3, activation='relu'),
    tf.keras.layers.MaxPooling2D(),
    tf.keras.layers.Conv2D(64, 3, activation='relu'),
    # tf.keras.layers.MaxPooling2D(),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(64, activation='relu'),
    tf.keras.layers.Dense(num_classes, activation='softmax')
])
```

Defining the model architecture in Google Colaboratory



Basic architecture of the custom neural network (CNN_Model_1)

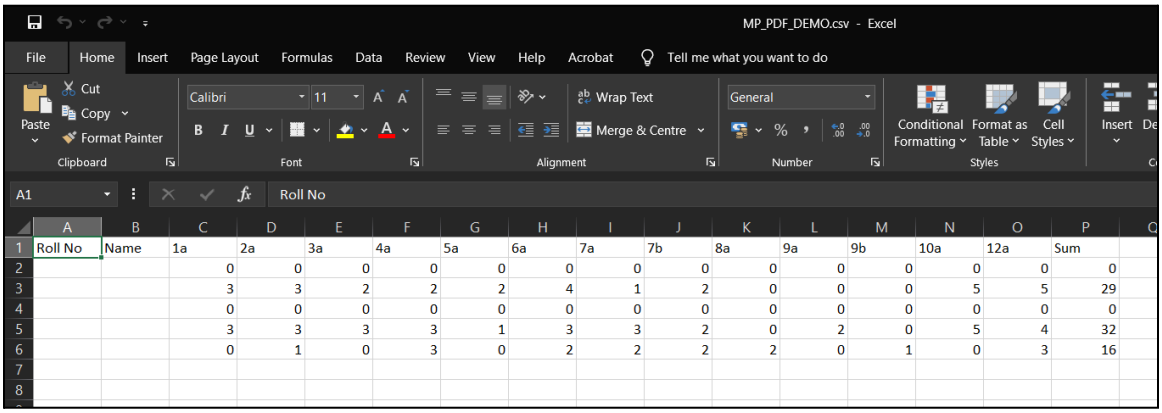
```
model.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(from_logits=True),
              metrics=['accuracy']) # compiling the model
```

```
history = model.fit(train_ds, validation_data=val_ds, epochs=30) # Fitting the model into training data
```

```
Epoch 2/30
532/532 [==============================] - 3s 6ms/step - loss: 0.2925 - accuracy: 0.9141 - val_loss: 0.2609 - val_accuracy: 0.9159
Epoch 3/30
532/532 [==============================] - 4s 7ms/step - loss: 0.1768 - accuracy: 0.9486 - val_loss: 0.2089 - val_accuracy: 0.9377
Epoch 4/30
532/532 [==============================] - 5s 9ms/step - loss: 0.1217 - accuracy: 0.9635 - val_loss: 0.1633 - val_accuracy: 0.9584
Epoch 5/30
532/532 [==============================] - 4s 7ms/step - loss: 0.1009 - accuracy: 0.9677 - val_loss: 0.1613 - val_accuracy: 0.9565
Epoch 6/30
532/532 [==============================] - 4s 8ms/step - loss: 0.0736 - accuracy: 0.9777 - val_loss: 0.1748 - val_accuracy: 0.9561
Epoch 7/30
532/532 [==============================] - 5s 9ms/step - loss: 0.0711 - accuracy: 0.9784 - val_loss: 0.1559 - val_accuracy: 0.9596
Epoch 8/30
532/532 [==============================] - 3s 6ms/step - loss: 0.0694 - accuracy: 0.9781 - val_loss: 0.1764 - val_accuracy: 0.9579
Epoch 9/30
532/532 [==============================] - 4s 7ms/step - loss: 0.0677 - accuracy: 0.9794 - val_loss: 0.1515 - val_accuracy: 0.9652
Epoch 10/30
532/532 [==============================] - 4s 7ms/step - loss: 0.0609 - accuracy: 0.9827 - val_loss: 0.1649 - val_accuracy: 0.9619
Epoch 11/30
532/532 [==============================] - 4s 7ms/step - loss: 0.0494 - accuracy: 0.9860 - val_loss: 0.1830 - val_accuracy: 0.9619
Epoch 12/30
532/532 [==============================] - 5s 9ms/step - loss: 0.0432 - accuracy: 0.9868 - val_loss: 0.1268 - val_accuracy: 0.9725
Epoch 13/30
532/532 [==============================] - 4s 7ms/step - loss: 0.0462 - accuracy: 0.9858 - val_loss: 0.1783 - val_accuracy: 0.9596
Epoch 14/30
532/532 [==============================] - 3s 6ms/step - loss: 0.0457 - accuracy: 0.9868 - val_loss: 0.2006 - val_accuracy: 0.9643
Epoch 15/30
532/532 [==============================] - 4s 8ms/step - loss: 0.0408 - accuracy: 0.9889 - val_loss: 0.3437 - val_accuracy: 0.9459
Epoch 16/30
532/532 [==============================] - 5s 9ms/step - loss: 0.0426 - accuracy: 0.9877 - val_loss: 0.1913 - val_accuracy: 0.9626
Epoch 17/30
532/532 [==============================] - 3s 6ms/step - loss: 0.0385 - accuracy: 0.9898 - val_loss: 0.1845 - val_accuracy: 0.9687
Epoch 18/30
532/532 [==============================] - 5s 9ms/step - loss: 0.0464 - accuracy: 0.9891 - val_loss: 0.2056 - val_accuracy: 0.9683
Epoch 19/30
532/532 [==============================] - 4s 8ms/step - loss: 0.0440 - accuracy: 0.9882 - val_loss: 0.1516 - val_accuracy: 0.9709
Epoch 20/30
532/532 [==============================] - 3s 6ms/step - loss: 0.0295 - accuracy: 0.9922 - val_loss: 0.1721 - val_accuracy: 0.9727
Epoch 21/30
532/532 [==============================] - 4s 8ms/step - loss: 0.0229 - accuracy: 0.9942 - val_loss: 0.1548 - val_accuracy: 0.9765
Epoch 22/30
532/532 [==============================] - 4s 8ms/step - loss: 0.0339 - accuracy: 0.9907 - val_loss: 0.1652 - val_accuracy: 0.9739
Epoch 23/30
532/532 [==============================] - 4s 7ms/step - loss: 0.0210 - accuracy: 0.9934 - val_loss: 0.1863 - val_accuracy: 0.9739
Epoch 24/30
532/532 [==============================] - 4s 7ms/step - loss: 0.0348 - accuracy: 0.9903 - val_loss: 0.1936 - val_accuracy: 0.9699
Epoch 25/30
532/532 [==============================] - 4s 8ms/step - loss: 0.0336 - accuracy: 0.9918 - val_loss: 0.2085 - val_accuracy: 0.9741
Epoch 26/30
532/532 [==============================] - 3s 6ms/step - loss: 0.0224 - accuracy: 0.9945 - val_loss: 0.1953 - val_accuracy: 0.9744
Epoch 27/30
532/532 [==============================] - 3s 6ms/step - loss: 0.0281 - accuracy: 0.9934 - val_loss: 0.1612 - val_accuracy: 0.9772
Epoch 28/30
532/532 [==============================] - 5s 9ms/step - loss: 0.0179 - accuracy: 0.9949 - val_loss: 0.1954 - val_accuracy: 0.9767
Epoch 29/30
532/532 [==============================] - 3s 6ms/step - loss: 0.0374 - accuracy: 0.9904 - val_loss: 0.2265 - val_accuracy: 0.9687
Epoch 30/30
532/532 [==============================] - 3s 6ms/step - loss: 0.0241 - accuracy: 0.9943 - val_loss: 0.1759 - val_accuracy: 0.9730
```

Snapshot of the training process

User interface of the system



MP_PDF_DEMO.csv - Excel

File Home Insert Page Layout Formulas Data Review View Help Acrobat Tell me what you want to do

Cut Copy Format Painter Paste Clipboard

Calibri 11 Font

Wrap Text Merge & Centre Alignment

General Number

Conditional Formatting Format as Table Cell Styles Styles

Insert

A1 Roll No

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Roll No	Name	1a	2a	3a	4a	5a	6a	7a	7b	8a	9a	9b	10a	12a	Sum
2			0	0	0	0	0	0	0	0	0	0	0	0	0	0
3			3	3	2	2	2	4	1	2	0	0	0	5	5	29
4			0	0	0	0	0	0	0	0	0	0	0	0	0	0
5			3	3	3	3	1	3	3	2	0	2	0	5	4	32
6			0	1	0	3	0	2	2	2	2	0	1	0	3	16
7																
8																

Final CSV output