

מטלת סיום קורס – למידת מכונה

במטלה זו החלטתי לחקור את הסיוג של תמונות אשר מרכיבות את הא'ב בשפת הסימנים. ברגע שיהיה לי מודל מספיק טוב, שייתן לי דיוק גבוה לסיוג התמונה בשפת סימנים לאות המקבילה באנגלית, זה יוכל לשמש ככלי עזר של תקשורת בין חרש לשומע. יהיה ניתן לפתח מעין אפליקציה/אתר שיוכנס אליה תמונה והיא "תתרגם" (תסווג) את התמונה לאות א'ב באנגלית.

קצת על הדאטה עליו אבצע את אימון המודל:
הדאטה מורכבת מ-34627 תמונות, class 25 כאשר כל תמונה מיוצגת על ידי פיקסלים- (פיצרים) $28 \times 28 = 784$.

במהלך הפרויקט רציתי לענות על שני שאלות מרכזיות:
(1) מציאת המודל עם הדיוק הטוב ביותר לסיוג האותיות
(2) לבדוק האם אני יכולה להוריד באופן משמעותי את מספר הפיקסלים עליהם אבצע את אימון המודל ועדיין לקבל אחוז דיוק גבוה – בחלק זה בחרתי לנסות על 2 אותיות.

חלק ראשון:

חלק זה בצעתי על 4 מודלים שונים:

	knn	svm	random forest	cnn
pixels 784	0.8103	0.782	0.814	0.971

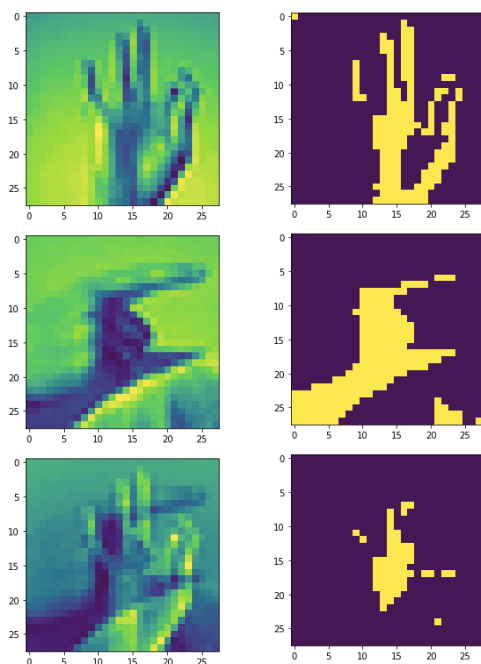
והמסקנה שלי בחלק זה הייתה :
המודל חחc הוא המודל בעל הדיוק הטוב ביותר עם דיוק של 0.971.
כאשר השכבות הן:

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 26, 26, 32)	320
max_pooling2d (MaxPooling2D)	(None, 13, 13, 32)	0
dropout (Dropout)	(None, 13, 13, 32)	0
conv2d_1 (Conv2D)	(None, 11, 11, 128)	36992
max_pooling2d_1 (MaxPooling2D)	(None, 6, 6, 128)	0
dropout_1 (Dropout)	(None, 6, 6, 128)	0
conv2d_2 (Conv2D)	(None, 4, 4, 512)	590336
max_pooling2d_2 (MaxPooling2D)	(None, 2, 2, 512)	0
dropout_2 (Dropout)	(None, 2, 2, 512)	0
flatten (Flatten)	(None, 2048)	0
dense (Dense)	(None, 4096)	8392704
dropout_3 (Dropout)	(None, 4096)	0
dense_1 (Dense)	(None, 1024)	4195328
dropout_4 (Dropout)	(None, 1024)	0
dense_2 (Dense)	(None, 256)	262400
dropout_5 (Dropout)	(None, 256)	0
dense_3 (Dense)	(None, 25)	6425

חלק שני:

בחרתי באופן אקראי class 2 עליהם ביצעתי את הבדיקה – באופן רנדומלי בחרתי שני סוגים – 6/5
חלק מהתהליך היה לנסות להבין איזה מהפיקסלים הן הפיקסלים המשמעותיים בכל תמונה – ולכן ביצעתי מספר פעולות כדי לזהותן:
כל פיקסל שערך שלו גבוה ממוצע הפיקסלים הכנסתי לו את הערך 1 אחרת 0.
ואזי ביצעתי xor על שני התמונות- כך שאם הערכים זהים תיתן 0 אחרת תיתן 1.

תמונה להמחשה:



לאחר מכן הרצתי את המודלים :

	knn	random forest	logistic regression
58 פיקסלים	0.663	0.715	0.601
784 פיקסלים	0.983	0.996	0.964

והמסקנה שלי בחלק זה הייתה :

(1) ההבדל של הרצת המודל על תמונה עם 58 פיקסלים או 748 לא משמעותית מבחינת זמן ריצה

(2) ההבדל של אחוז הדיוק היה משמעותי עם מספר פיקסלים גבוה יותר.