

תיאור כללי של הפרויקט:

- מימוש PIPE לזיהוי אובייקטים.
- אימון ולכונן את ה-PIPE על תת קבוצה מסט הנתונים Caletch101.
- בחינת ביצועי המסווג על תת קבוצה אחרת מסט הנתונים.
- ניתוח התוצאות.

סט הנתונים:

- סט הנתונים כולל 101 מחלקות שבכל אחת 31-800 תמונות.
- טענו את התמונות, העברנו אותן ל-gray scale ושינינו את הגודל שלהם בהתאם להיפר פרמטר S.
- חילוק הנתונים
 - דיבגנו וכיוונו את ה-PIPE על עשרת המחלקות הראשונות.
 - הרצנו את האלגוריתם עם ההיפר-פרמטרים הטובים ביותר שנמצאו על עשרת המחלקות השניות (מחלקות 11-20, Fold 2).
- אימנו ובחנו את המודל: אימון על 20 התמונות הראשונות בכל מחלקה ובחינה על 20 תמונות אחרות (אם לא היו 40 תמונות אז נשארו פחות תמונות לאימון).

ה-PIPE שמומש:

ה-PIPE שבחרנו לממש הוא SVM+HOG. מסווג ה-SVM אומן באמצעות שני סוגים של kernels-לינארי וגאוסיאני (RBF).

כיוון ההיפר פרמטרים:

כיוונו את ההיפר פרמטרים לכל סוג של kernel בנפרד. בדו"ח אנחנו מציגים את ה-kernel הלינארי-המודל שהביא לתוצאות דיוק הטובות ביותר. (כיוון הפרמטרים על המודל השני נמצא בנספחים).

ההיפר פרמטרים שבחרנו לכוון הם:

- S-גודל התמונה – אחרי שינוי של התמונה המקורית בסט הנתונים.
- Pixels-Per-Cells-גודל תא מרחבי - פרמטר קלט לאלגוריתם HOG. גודל התא בו מחושב השיפוע.
- Num-Of-Bins-מספר ה"בינים" של הכיוונים. פרמטר קלט לאלגוריתם HOG. מספר הכיוונים האפשריים של הגרדיאנט.
- Block norm method-קלט לאלגוריתם ה-HOG. שיטת הנורמליזציה עבור כל בלוק.
- C-פרמטר של אלגוריתם ה-SVM. אחראי על ה-tradeoff בין סיבוכיות להתאמת המודל.
- RBF kernel-Gamma. כאשר $\gamma = -\frac{1}{2\sigma^2}$

תהליך הכוון היה איטרטיבי ונמשך עד התכנסות. נתאר כעת את התהליך עבור המודל הלינארי:

תחילה בחרנו לכוון את הפרמטרים C ו-S באמצעות Grid Search (הגדרת מרחב חיפוש ומציאת הפתרון האופטימלי במרחב הנתון), כאשר הפרמטרים האחרים מקובעים לערכים נפוצים אשר מצאנו באינטרנט. לאחר מציאת הערכים האופטימליים במרחב החיפוש של C ו-S, בחרנו לכוון את הפרמטרים Pixels-Per-Cells ו-Num-Of-Bins באמצעות Grid Search. השלב האחרון היה לכוון את הפרמטרים pixels_per_block ו-block norm method. בחרנו לעשות זאת בנפרד בגלל זמן ריצה ארוך.

זוהי הייתה האיטרציה הראשונה של התהליך. האיטרציה השנייה התחילה מחדש במציאת פרמטרים C ו-S רק הפעם עם 4 ערכים לפרמטרים האחרים שמצאנו באיטרציה הקודמת (ולא לפי מה שמצאנו באינטרנט). המשך האיטרציה הוא זהה והתהליך נמשך עד התכנסות – כלומר כאשר אין שינוי באף אחד מהערכים של ההיפר פרמטרים.

חשוב לציין, את תהליך הולידציה שעשינו על מנת לבדוק את טיב המודל. בחרנו לממש K-Folds לולידציה. חילקנו את סט האימון ל-5 folds, כל אחד מהם כולל 4 תמונות מכל מחלקה (בסך הכל 40 תמונות בכל fold).

***הערה:** עבור המודל הגאוסיאני (kernel='rbf') האיטרציה התחילה במציאת ערכי gamma ע"י grid-search, כאשר שאר הערכים מקבועים לפרמטרים שנמצאו עבור המודל הלינארי. לאחר מציאת ה-gamma האופטימלי במרחב האיטרציה נמשכה בדיוק כפי שתואר לעיל עבור המודל הלינארי, עד התכנסות של הפרמטרים.

הטבלה להלן מראה את היפר הפרמטרים, את טווחי החיפוש, את הערך הנבחר ואת אחוז השגיאה הממוצע על 5 ה-Folds (סט הולידציה).

Iteration	Hyper Parameter	Stage	Range	Jumps	Selected Value	Error Rate
1	C	1	10^-3, 10^2	exp	8-15	0.325
		2	1:100	10		
		3	8:15	1		
	Image size (S)	1	100,300,600,800	X	83	
		2	80:120	5		
		3	79:90	2		
	Number of oriented gradient (num_of_bins)	1	1:30	10		0.285
		2	5:15	1	13	
	Spatial cell size (Pixels per cell)	1	1:30	10		
		2	5:15	1	13	
	Cells per block	1	1:5	1	2	0.285
	Block norm method	1	"L1", "L2", "L1-Sqrt", "L2-Hys"	X	L2-Hys	0.285
2	C	1	10^-3, 10^3	exp	1-1000	0.28
		2	1,10,50,100,200,500,1000	X		
		3	1:1000	3		
	Image size (S)	1	75:100	2	81,93	
		2	80,81,82,92,93,94	X		
		3	81,93	X		

	Number of oriented gradient (num_of_bins)	1	1:30	5	13	0.28
		2	8:16	1		
	Spatial cell size (Pixels per cell)	1	1:30	5	13	
		2	8:16	1		
	Cells per block	1	1:5	1	2	0.28
	Block norm method	1	"L1", "L2", "L1-Sqrt", "L2-Hys"	X	L2-Hys	0.28

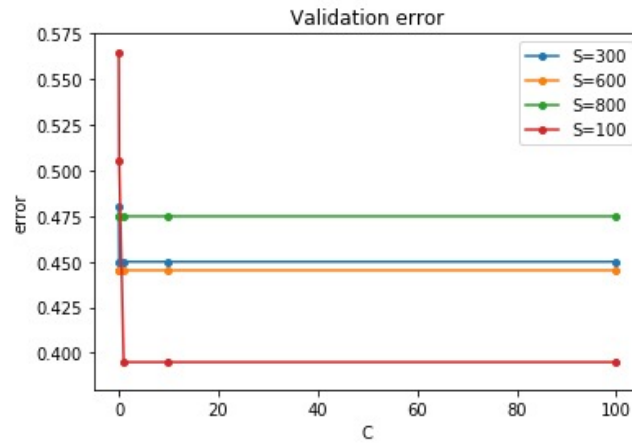
גרפים של כוונן ההיפר פרמטרים:

- איטרציה מספר 1

○ כיוון C ו-S

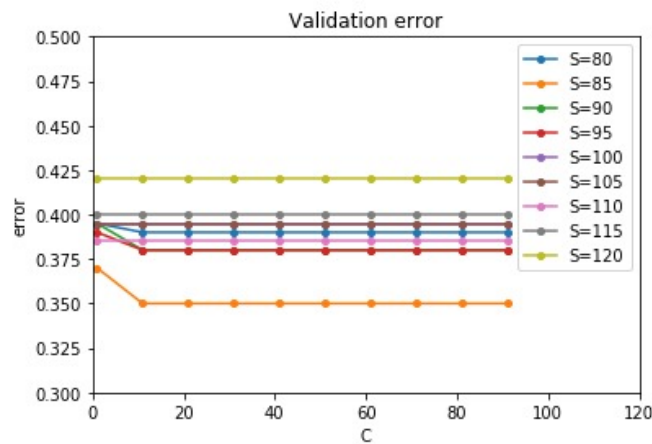
▪ שלב 1- הבנת סדרי הגודל של C ו-S:

$$C \in \{10^i | i = -2, -1, 0, 1, 2\}, S \in \{100, 300, 600, 800\}$$



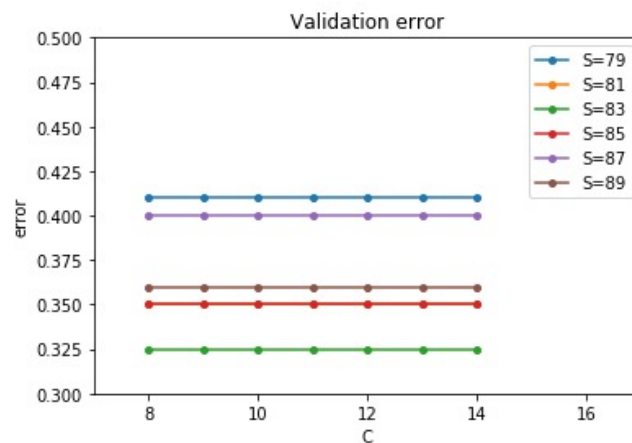
▪ שלב 2 – צמצום מרחב הערכים של C ו-S

$$C \in \{1 + 10 * i | i = 0, 1, \dots, 9\}, S \in \{80, 85, 90, 95, 100, 105, 110, 115\}$$



▪ שלב 3 – מציאת ערך אופטימלי במרחב חיפוש:

$$C \in \{8, 9, \dots, 15\}, S \in \{79, 81, 83, 85, 87, 89\}$$

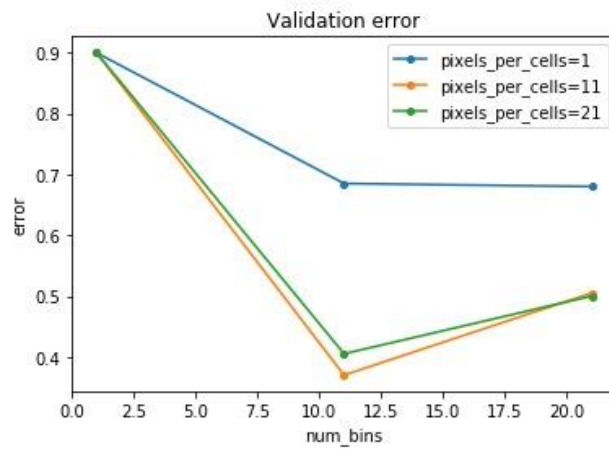


נבחרו $C = 8: 15, S = 83$

כיוון num_of_bins ו- pixels_per_cells :

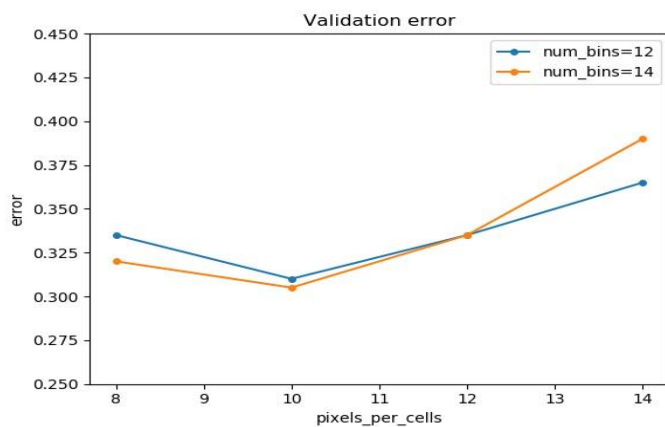
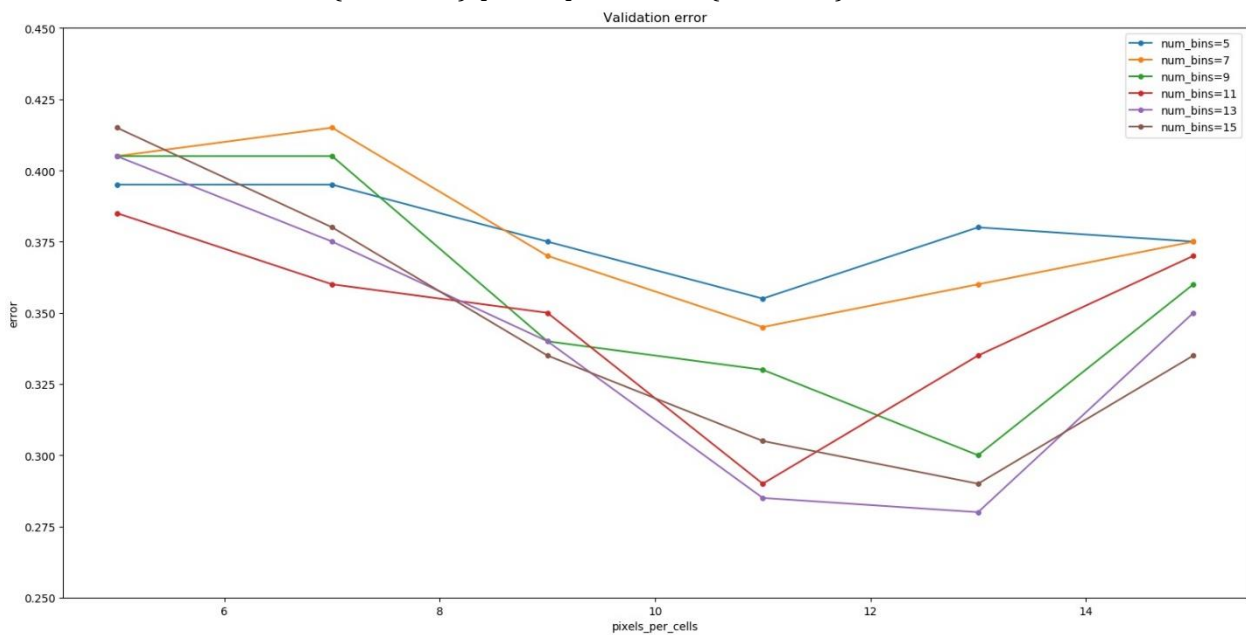
■ שלב 1- הבנת סדרי הגודל

$\text{num_of_bins} \in \{1, 11, 21\}$, $\text{pixels_per_cells} \in \{1, 11, 21\}$



■ שלב 2 – צמצום ומציאת הערך האופטימלי:

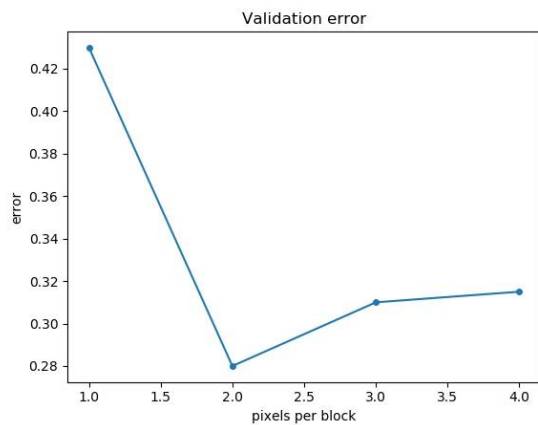
$\text{num_of_bins} \in \{5, 6, \dots, 15\}$, $\text{pixels_per_cells} \in \{5, 6, \dots, 15\}$



נבחרו $\text{numBins} = 13$, $\text{pixelsPerCells} = 13$

כיוון `pixels_per_block`:

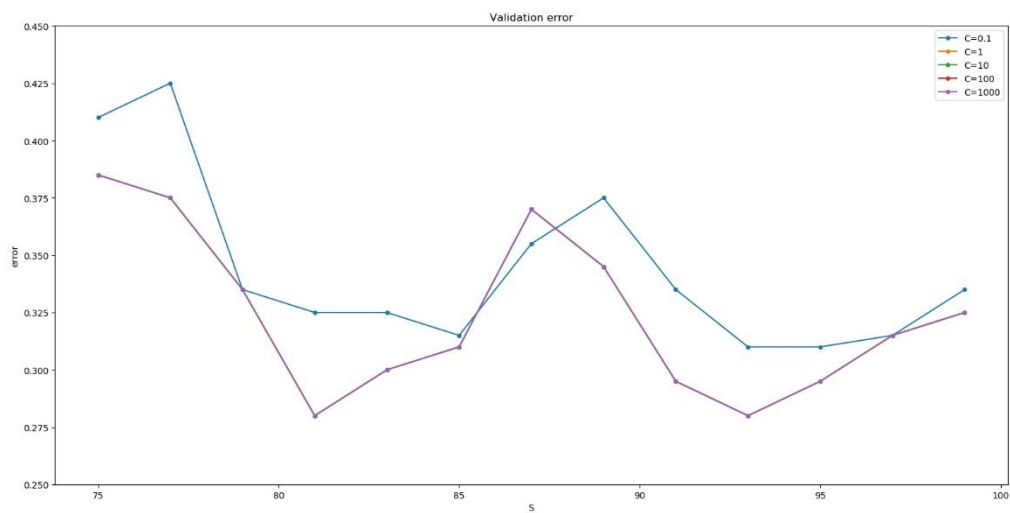
נבחר $pixelsPerBlock = 2$



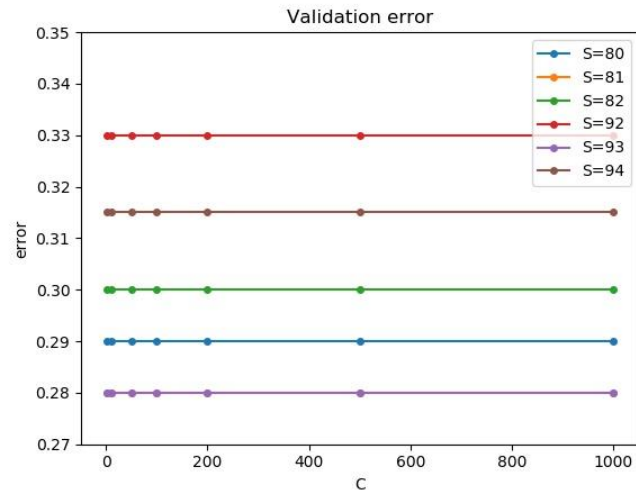
איטרציה מספר 2:

כיוון `C` ו-`S`

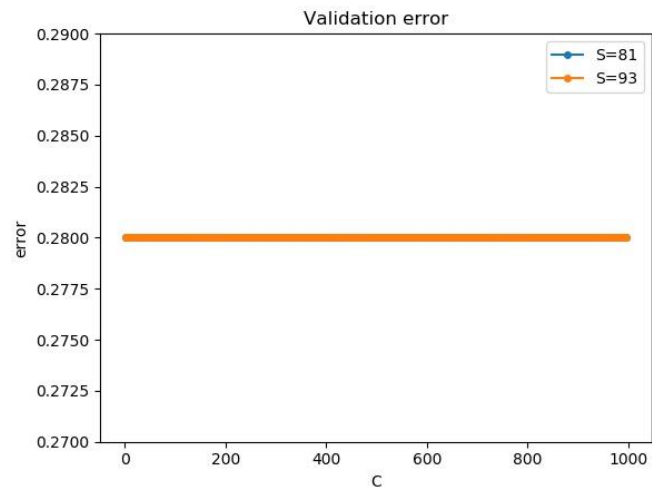
- שלב 1 – צמצום מרחב החיפוש
 $C \in \{10^i | i = -2, -1, 0, 1, 2\}, S \in \{75, 77, \dots, 99\}$



- שלב 2 – צמצום נוסף של מרחב החיפוש
 $C \in \{1, 10, 50, 100, 200, 500, 1000\}, S \in \{80, 81, 82, 92, 93, 94\}$



■ שלב 3 – מציאת ערך אופטימלי
 $C \in \{1, 4, \dots, 1000\}, S \in \{81, 93\}$



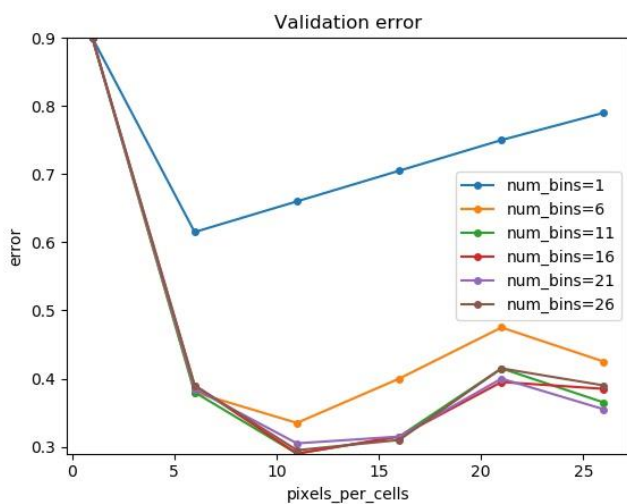
נבחרו $C = 10, S = 93$

כיוון num_of_bins ו- pixels_per_cells :

■ שלב 1 – צמצום מרחב החיפוש

$\text{num_of_bins} \in \{1, 6, 11, 16, 21, 26\},$

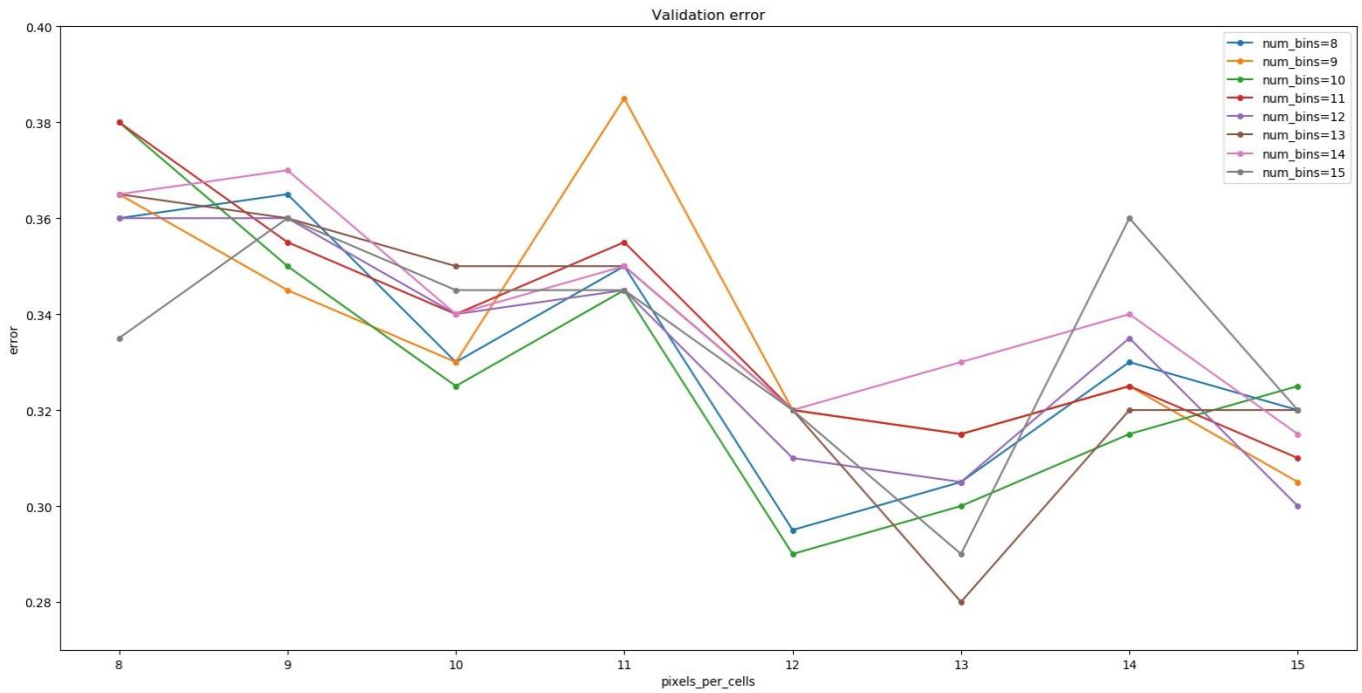
$\text{pixels_per_cells} \in \{1, 6, 11, 16, 21, 26\}$



■ שלב 2 – מציאת ערך אופטימלי

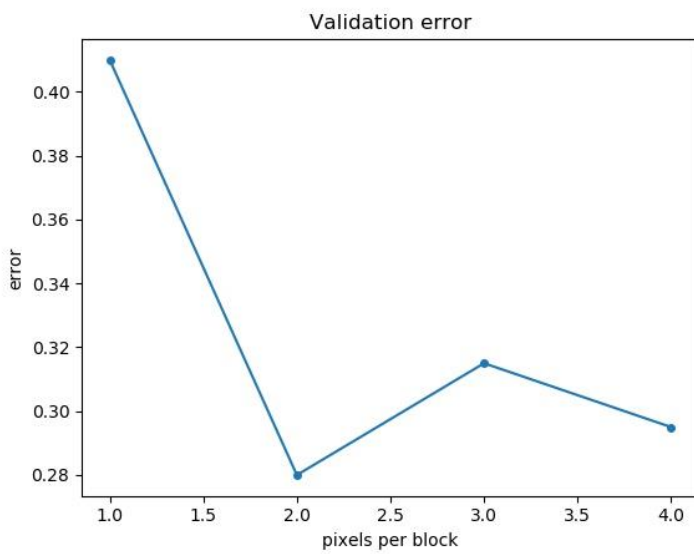
$\text{num_of_bins} \in \{8, 9, \dots, 15\}$,

$\text{pixels_per_cells} \in \{8, 9, \dots, 15\}$



נבחרו $\text{numBins} = 13, \text{pixelsPerCells} = 13$

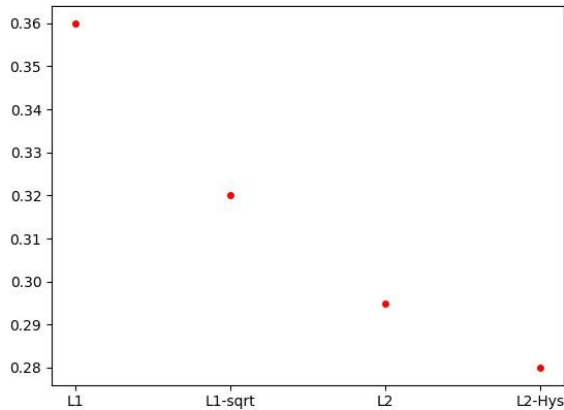
כיוון pixels_per_block :



נבחר $\text{pixelsPerCells} = 2$

כיוון block_norm:

- שלב 1 – מציאת ערך אופטימלי
 $BlockNorm \in \{L1', L2', L1 - Sqrt', L2 - Hys'\}$



נבחר $BlockNorm = L2 - Hys$

תוצאות הבחינה:

אחוזי השגיאה על סט האימון (עבור Fold 2) הוא 32.5%











מטריצת המבוכה:



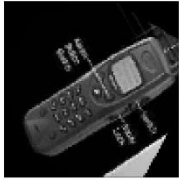




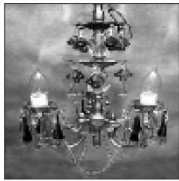
מטריצת המבוכה היא מטריצה בגודל $M \times M$ כאשר M מציין את מספר המחלקות. התא M_{ij} מציין את מספר הפעמים שדוגמא מהמחלקה i סווגה למחלקה j .

תוויות שסווגו											תוויות אמיתיות
20	19	18	17	16	15	14	13	12	11		
0	3	0	0	0	1	0	1	1	14	11	
2	0	0	1	0	2	0	0	14	1	12	
0	0	2	2	0	5	0	10	0	1	13	
1	0	0	0	0	2	14	1	1	1	14	
2	0	0	1	1	11	0	5	0	0	15	
0	0	0	0	20	0	0	0	0	0	16	
2	0	0	11	0	4	0	0	0	3	17	
0	0	18	0	0	1	1	0	0	0	18	
1	12	0	1	1	5	0	0	0	0	19	
11	0	0	1	0	4	1	1	2	0	20	

שגיאות ויזואליזציה:

לכל מחלקה אנו מציגים תמונות של שתי השגיאות הגדולות ביותר בתמונות של המחלקה (כלומר תמונות מהמחלקה שסווגו לא נכון).

מחלקה	תמונה 1	תמונה 2
brontosaurus		
Buddha		
Butterfly		
Camera		
Cannon		

		Ceiling fan
		Cellphone
		Chair
		Chandelier

(kernel='rbf') כיוון פרמטרים עבור המודל הגאוסיאני

Iteration	Stage	Hyper Parameter	Range	Jumps	Selected Value	Error Rate
1	1	gamma	0.01,0.05,0.1,0.5,1,5,10,50	X	0.45	0.35
	2		0.1:1	0.05		
	3		0.4:0.5	0.01		
	1	C	10^-3, 10^2	exp	10	0.345
	2		5:100	10		
	3		10	X		
	1	Image size (S)	100,300,600,800	X	96	
	2		80:110	5		
	3		90:100	1		
	1	Number of oriented gradient (num_of_bins)	1:30	5	7	0.33
	2		4:10	1		
	1	Spatial cell size (Pixels per cell)	1:25	5	16	
	2		12:20	1		
	1	Cells per block	1:5	1	2	0.33
	1	Block norm method	"L1", "L2", "L1-Sqrt", "L2-Hys"	X	L1	0.33
2	1	gamma	0.05:1	0.05	0.46	0.325
	2		0.4:0.6	0.01		
	1	C	10:50	10	20	0.32
	2		10:30	2		
	3		20	X		
	1	Image size (S)	70:110	10	98	
	2		80:100	2		
	3		97:99	1		

	1	Number of oriented gradient (num_of_bins)	4:10	1	8	0.32
	1	Spatial cell size (Pixels per cell)	12:20	1	16	
	1	Cells per block	1:5	1	2	0.32
	1	Block norm method	"L1", "L2", "L1-Sqrt", "L2-Hys"	X	L1	0.32
3	1	gamma	0.05:1	0.05	0.96	0.295
	2		0.85:1.45	0.05		
	3		0.9:1	0.01		
	1	C	10:28	2	20	0.295
	1	Image size (S)	90:108	2	98	
	1	Number of oriented gradient (num_of_bins)	4:10	1	8	0.295
	1	Spatial cell size (Pixels per cell)	12:20	1	16	
	1	Cells per block	1:5	1	2	0.295
	1	Block norm method	"L1", "L2", "L1-Sqrt", "L2-Hys"	X	L1	0.295

גרפים של כוונון ההיפר פרמטרים:

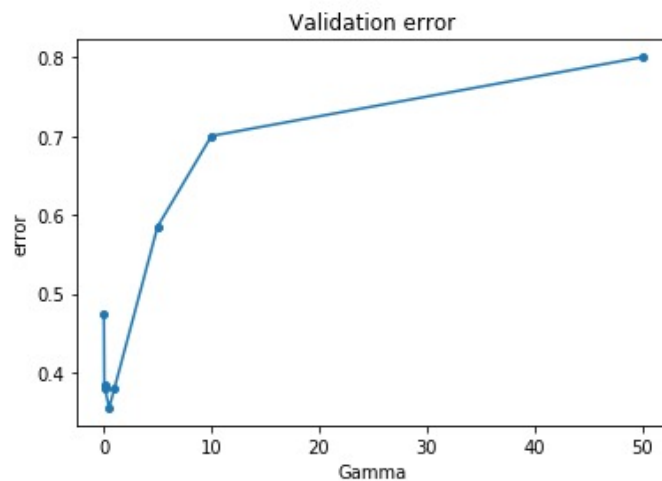
הגרפים להלן מציגים את אחוזי השגיאה הממוצעת על סט הולידציה כפונקציה של הערך של ההיפר פרמטרים:

איטרציה מספר 1:

- כיוון gamma

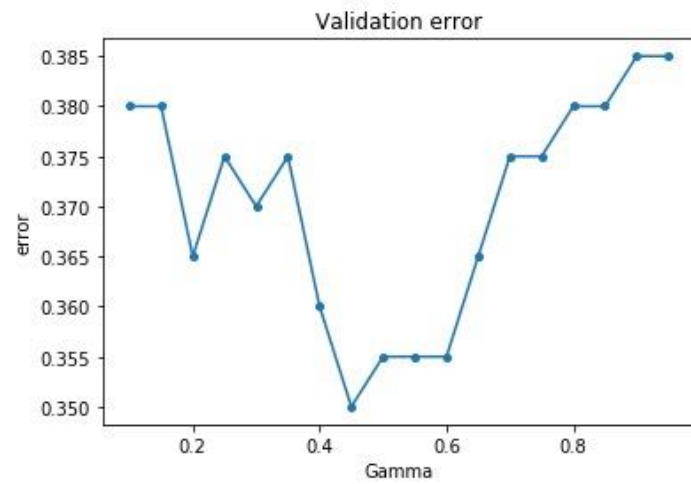
שלב 1- הבנת סדר הגודל של gamma:

$$\gamma \in \{0.01, 0.05, 0.1, 0.5, 1, 5, 10, 50\}$$



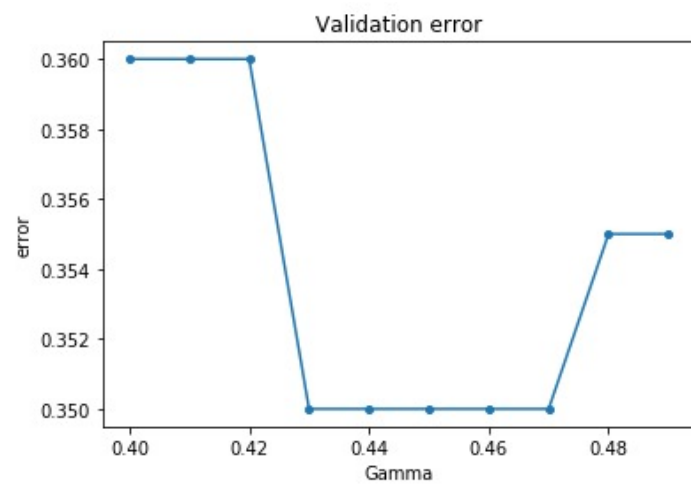
שלב 2- צמצום מרחב החיפוש:

$$\gamma \in \{0.1 + 0.05 * i | i = 0, 1, \dots, 19\}$$



שלב 3 – מציאת ערך אופטימלי:

$$\gamma \in \{0.4 + 0.01 * i | i = 0, 1, \dots, 9\}$$

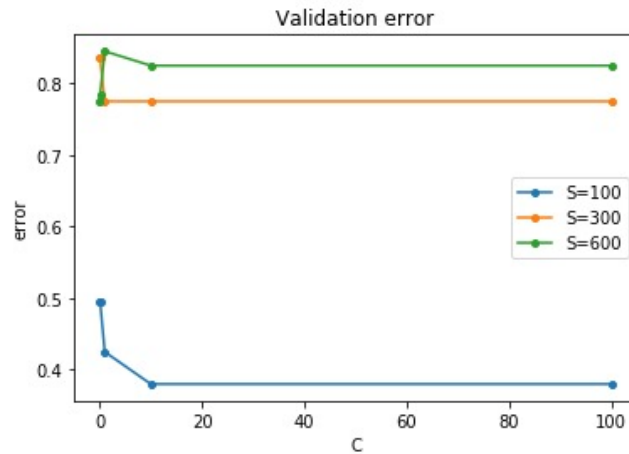


ערך נבחר $\gamma = 0.45$

- כיוון C ו-S

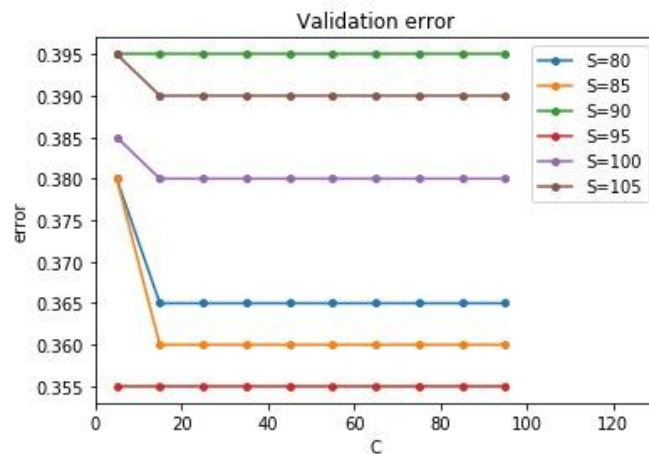
שלב 1- הבנת סדרי הגודל של C ו-S:

$$C \in \{10^i | i = -2, -1, 0, 1, 2\}, S \in \{100, 300, 600\}$$



שלב 2 – צמצום מרחב הערכים:

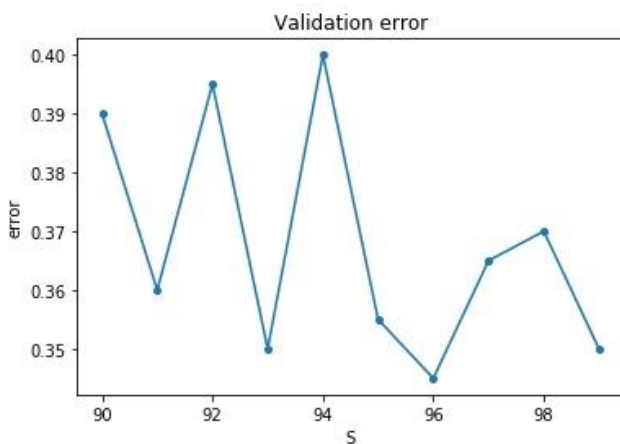
$$C \in \{5 + 10 * i | i = 0, \dots, 9\}, S \in \{50 + 5 * i | i = 0, 1, \dots, 19\}$$



שמנו לב שערך C לא משנה לערכי המודל ולכן קיבענו $C = 10$ לשלב 3.

שלב 3 – מציאת ערך אופטימלי:

$$C = 10, S \in \{90, 91, \dots, 99\}$$

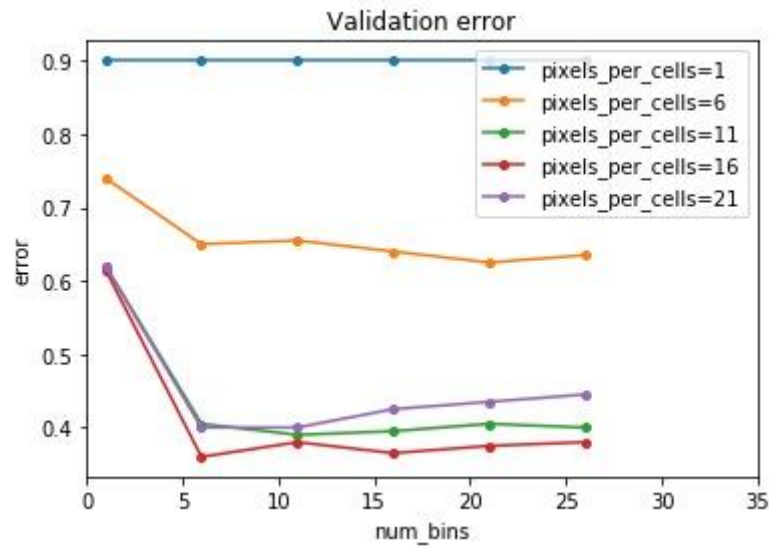


נבחר $C = 10, S = 96$

• כיוון num_of_bins ו-pixels_per_cells:

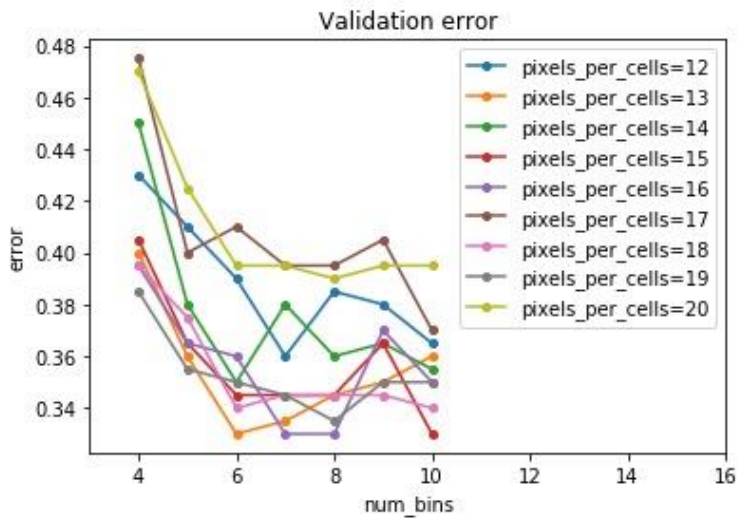
שלב 1- הבנת סדרי הגודל

$\text{num_of_bins} \in \{1,6,11,16,21,26\}$,
 $\text{pixels_per_cells} \in \{1,6,11,16,21\}$



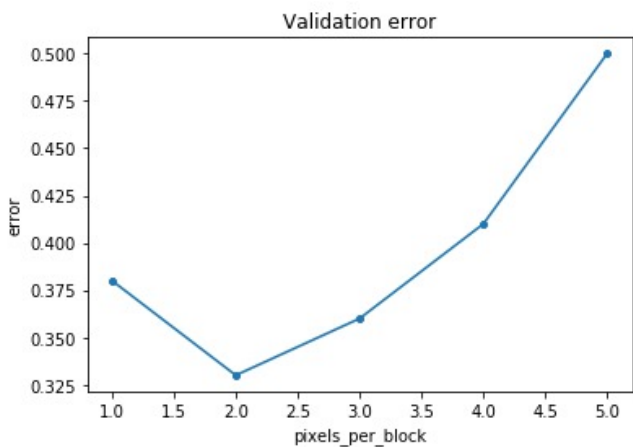
שלב 2- צמצום מרחב החיפוש ומציאת הערך האופטימלי

$\text{num_of_bins} \in \{4,5,\dots,10\}$, $\text{pixels_per_cells} \in \{12,13,\dots,20\}$



נבחר, $\text{numBins} = 7$

$\text{pixelsPerCells} = 16$



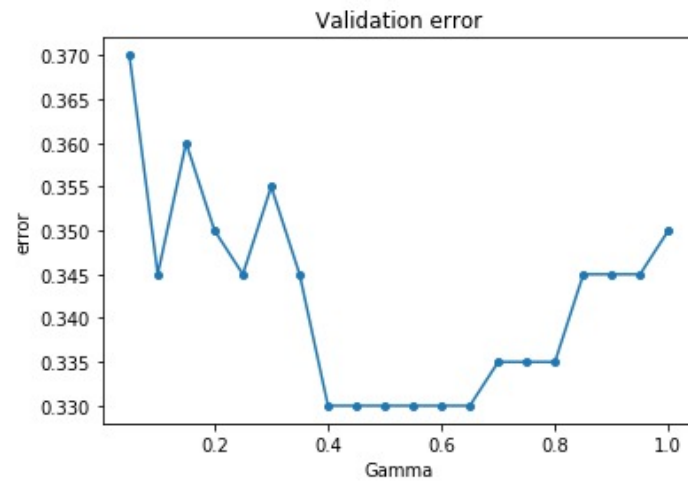
• כיוון pixels_per_blocks :

נבחר $\text{pixelsPerBlock} = 2$

איטרציה 2:

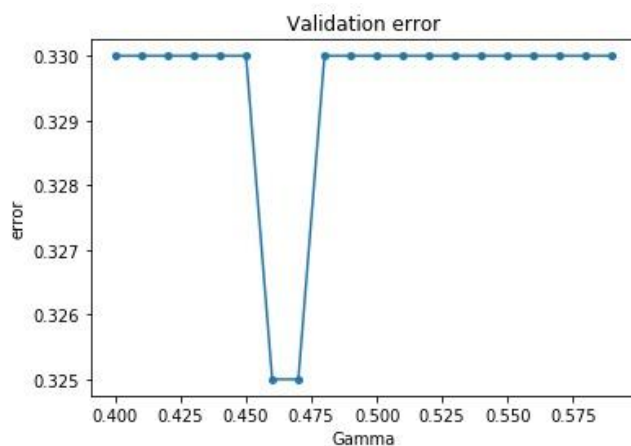
- כיוון gamma:

שלב 1 – צמצום מרחב הערכים (הפעם התחלנו עם מרחב ערכים קטן יותר)
 $\gamma \in \{0.05 * i | i = 1, \dots, 20\}$



שלב 2 – מציאת ערך אופטימלי

$\gamma \in \{0.4 + 0.01 * i | i = 0, 1, \dots, 19\}$

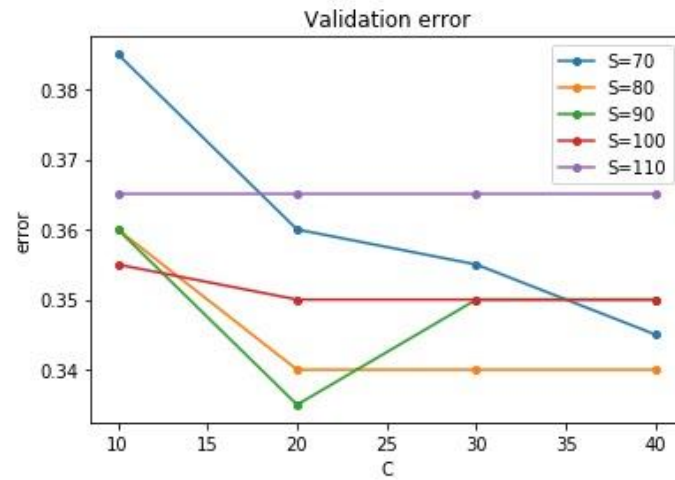


ערך נבחר $\gamma = 0.46$

- כיוון C ו-S:

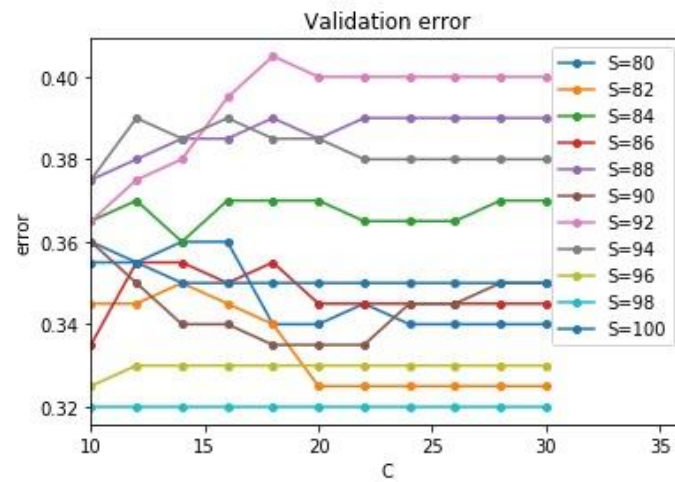
שלב 1 – צמצום מרחב הערכים

$$C \in \{10 * i | i = 1, \dots, 5\}, S \in \{70 + 10 * i | i = 0, 1, \dots, 4\}$$



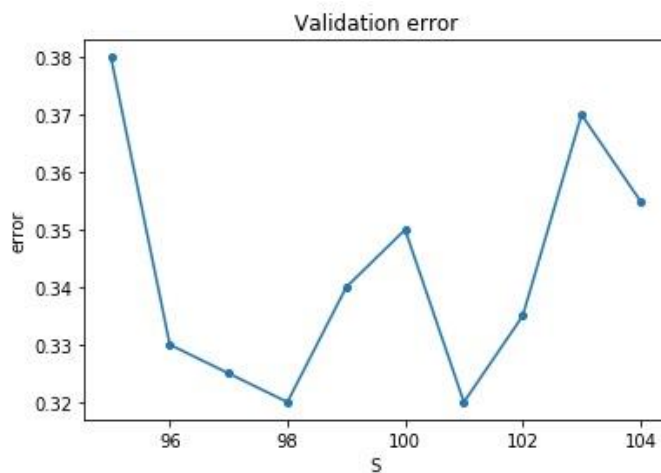
שלב 2 – ציליה פנימה במרחב הערכים

$$C \in \{10 + 2 * i | i = 0, \dots, 10\}, S \in \{80 + 2 * i | i = 0, 1, \dots, 10\}$$



שלב 3 – מציאת ערך אופטימלי

$$C = 20, S \in \{95, \dots, 104\}$$

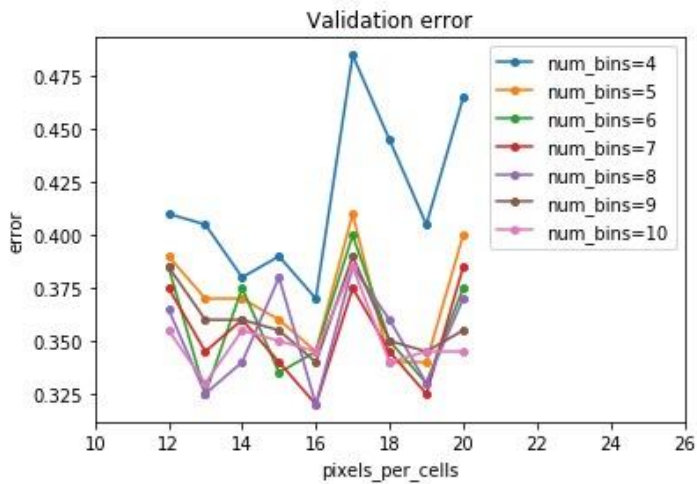


נבחר $C = 20, S = 98$

• כיוון num_of_bins ו-pixels_per_cells

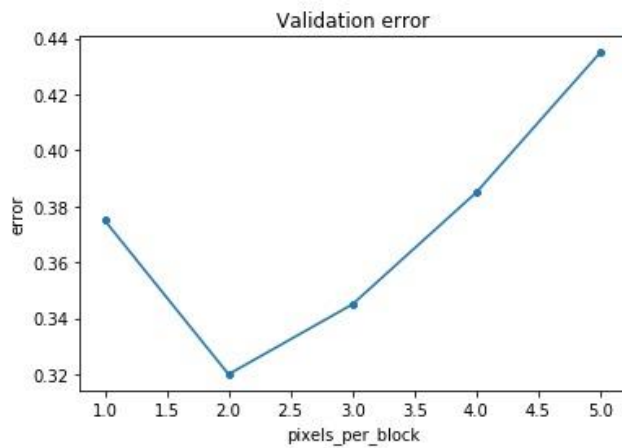
שלב 1 - מציאת ערך אופטימלי

$\text{num_of_bins} \in \{4, \dots, 10\}$, $\text{pixels_per_cells} \in \{12, \dots, 20\}$



ערכים נבחרים:

$\text{numBins} = 8, \text{pixelsPerCells} = 16$



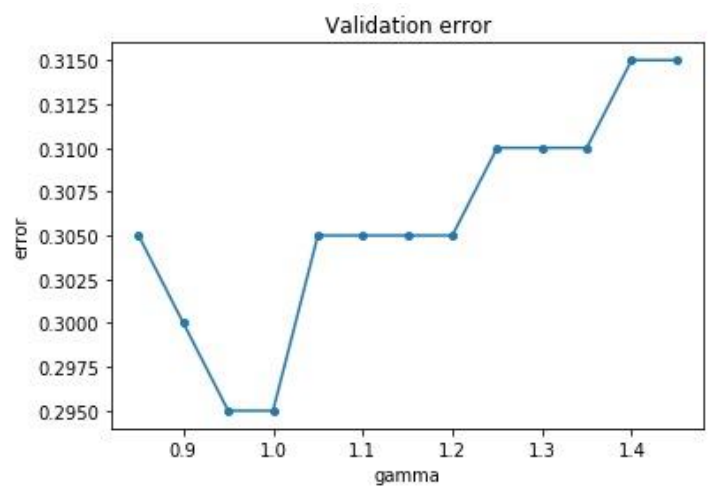
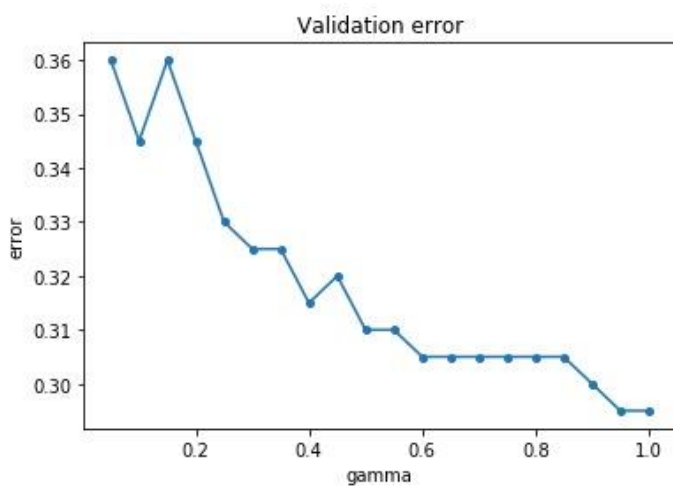
• כיוון pixels_per_block:

ערך נבחר: $\text{pixelsPerCells} = 16$

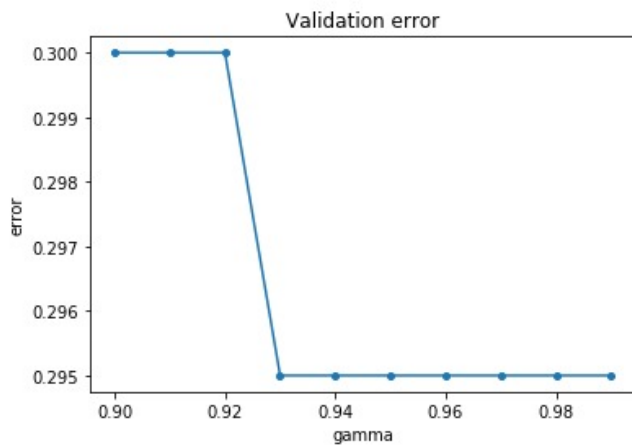
איטרציה 3:

• כיוון gamma:

שלב 1 - הבנת סדר הגודל



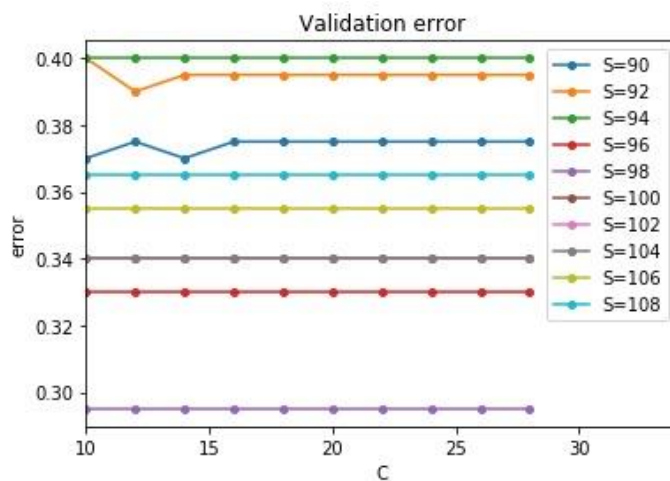
שלב 2 – מציאת ערך אופטימלי



נבחר $\gamma = 0.96$

- כיוון C ו-S:

שלב 1 – מציאת ערך אופטימלי



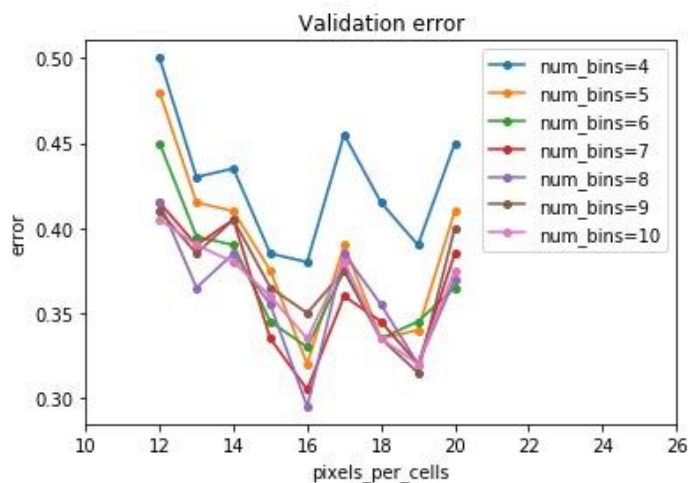
נבחר:

$$C = 20, S = 98$$

- כיוון num_of_bins ו-pixels_per_cells:

שלב 1 – מציאת ערך אופטימלי

$$\text{num_of_bins} \in \{4, \dots, 10\}, \quad \text{pixels_per_cells} \in \{12, \dots, 20\}$$



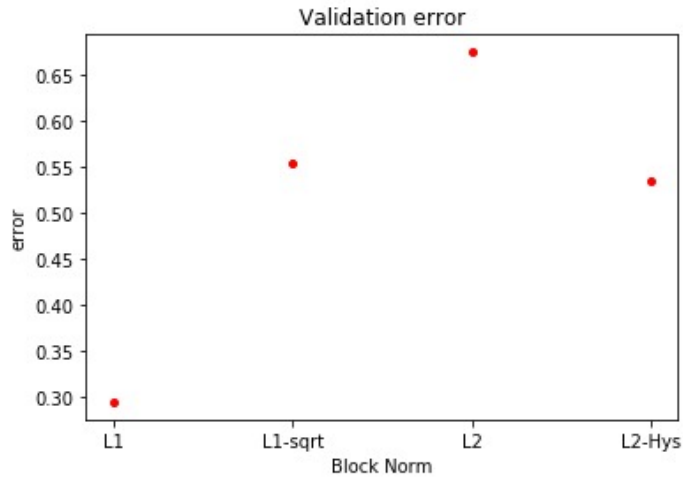
נבחר $\text{numBins} = 8$,

$\text{pixelsPerCells} = 16$

• כיוון block_norm:

שלב 1 – מציאת ערך אופטימלי

$$BlockNorm \in \{L1', L2', L1 - Sqrt', L2 - Hys'\}$$



נבחר

$$BlockNorm = L1'$$

לאחר האיטרציה הזו הפסיקו להיות שינויים בערכי הפרמטרים ולכן התהליך נעצר. אחוז השגיאה עמד על 29.5% ולכן המודל שנבחר הוא המודל הלינארי לו אחוזי שגיאה נמוכים יותר (28).