22/03/2023

New GT

Compute centroids and intra/inter dists. for current GT classes

- Intra dist for filling (0.07) > Intra dist for not_filling_peso (0.01)
- filling 4x4 (0.16) > not filling peso 4x4 (0.068)
- not filling intra dist for 2x2 (0.01) < 4x4 (0.068)

Inter/Intra distances (not filling)

- Intra (same class): Max distance from centroid (sample mas alejado)
- Inter (other class): Min distance from centroid (sample from other class that is closest to centroid →Centroid A to sample from B is different that centroid B from sample A → Better to measure distance between centroids? Then there is no min value

2x2	Class A	Class B	Class C	Class D
Class_A	0.001	0.001	0.001	0.007
Class_B	0	0.011	0.001	0.004
Class_C	0.001	0.001	0.004	0.006
Class_D	0.003	0.002	0.002	0.011

5x5	Class A	Class B	Class C	Class D
Class_A	1.973	1.82	1.783	1.484
Class_B	0.196	2.895	1.257	2.192
Class_C	1.428	1.515	2.4	2.587
Class_D	1.979	2.334	2.342	1.393

4x4	Class A	Class B	Class C	Class D
Class_A	0.954	0.879	1.024	1.626
Class_B	0.056	2.142	0.899	1.429
Class_C	0.063	0.893	1.508	1.725
Class_D	1.95	1.826	1.759	0.069

10x10	Class A	Class B	Class C	Class D
Class_A	3.49	4.767	4.719	4.077
Class_B	1.867	5.274	2.521	5.094
Class_C	4.833	2.736	4.299	5.618
Class_D	4.874	5.012	5.261	2.435

Intra should be less than inters!

New GT

- Remove samples that make intra larger → Move them where?
- Remove samples that make inter smaller → Move them where?
- Compute intra/inter again

Changed success extraction method - not filling peso

Combining class label assignment and selecting the one that maximises total success

- n_clusters=4, init=**'random'**, n_init=20, max_iter=500, tol=1e-8, random_state=0

	2x2	3x3	4x4	5x5	6x6	7x7	8x8	9x9	10x10
Success_A	0	100	0	92.857	100	100	100	100	96.429
Success_B	73.333	15.556	57.778	40	0	57.778	0	62.222	55.556
Success_C	45.455	36.364	27.273	0	27.273	27.273	63.636	0	27.273
Success_D	33.333	100	100	100	0	100	0	0	100
TOTAL	44.444	50	38.889	55.556	34.444	70	38.889	62.222	67.778

- n_clusters=4, init=**k-means++**, n_init=20, max_iter=500, tol=1e-8, random_state=0

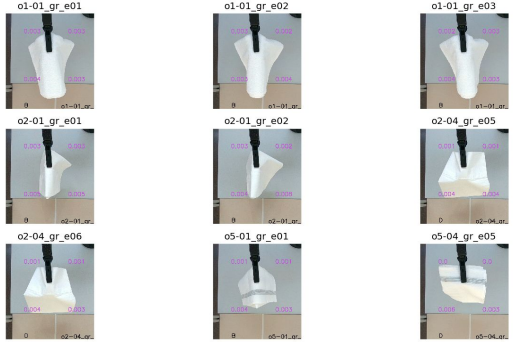
	2x2	3x3	4x4	5x5	6x6	7x7	8x8	9x9	10x10
Success_A	100	100	100	92.857	100	100	100	100	0
Success_B	20	20	26.667	44.444	60	26.667	26.667	26.667	62.222
Success_C	0	36.364	0	0	0	54.545	72.727	72.727	0
Success_D	33.333	100	100	100	0	100	100	100	100
TOTAL	43.333	52.222	51.111	57.778	61.111	57.778	60	60	37.778

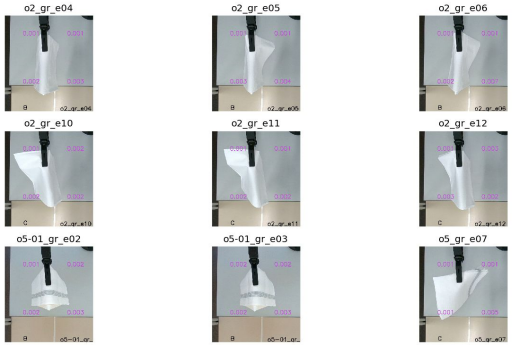
- n_clusters=4, **init=[init_centr]** , n_init=20, max_iter=500, tol=1e-8, **random_state=None** -> samples with LESS intra dist to centroid of class

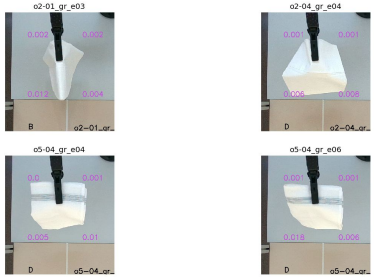
resulting clusters in next slides

	2x2	3x3	4x4	5x5	6x6	7x7	8x8	9x9	10x10
Success_A	100	100	100	96.429	100	100	100	100	100
Success_B	13.333	11.111	33.333	37.778	62.222	82.222	88.889	55.556	55.556
Success_C	36.364	36.364	36.364	45.455	63.636	27.273	63.636	27.273	27.273
Success_D	50	100	100	100	100	100	100	100	100
TOTAL	45.556	47.778	58.889	61.111	76.667	82.222	90	68.889	68.889

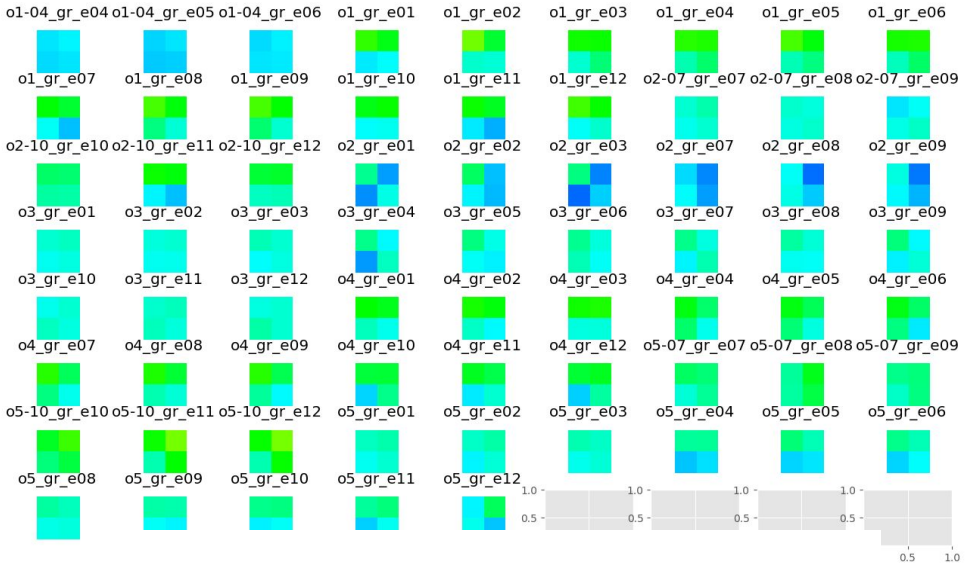
not filling peso 2x2

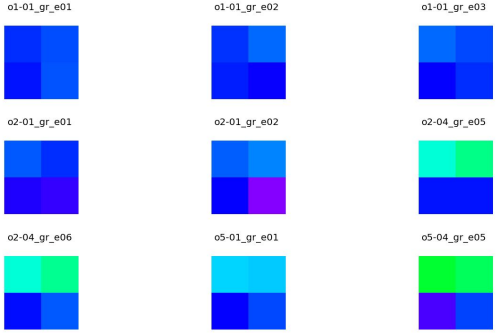


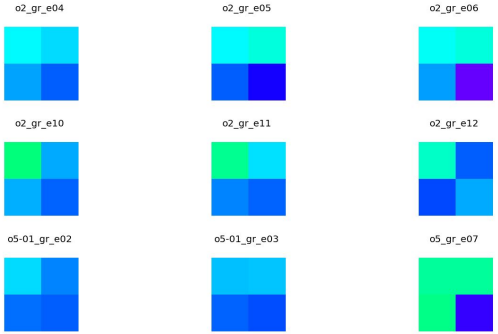


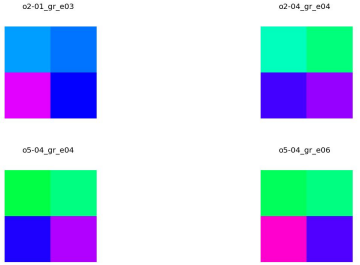


not filling peso 2x2

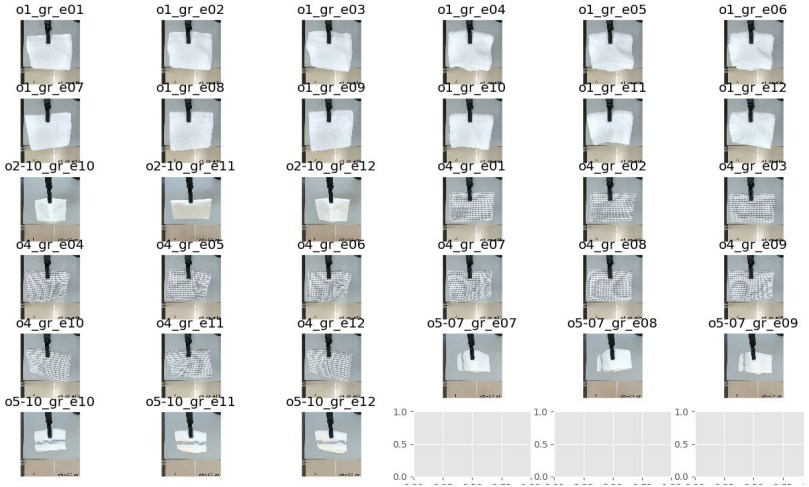


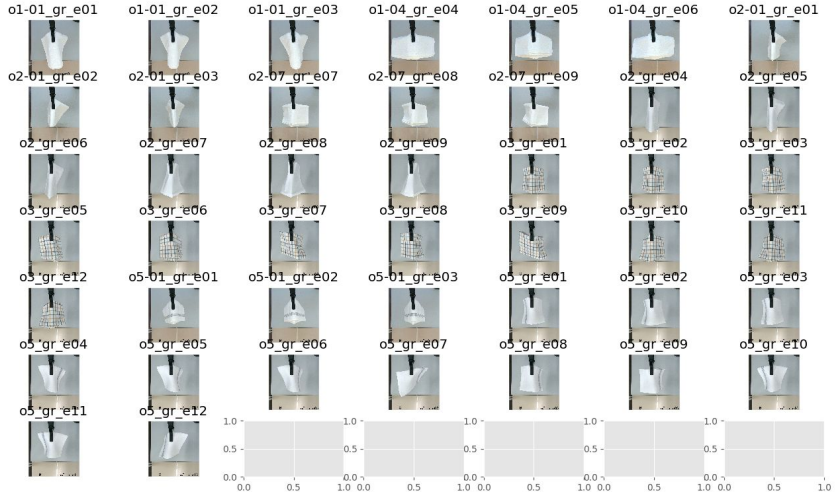


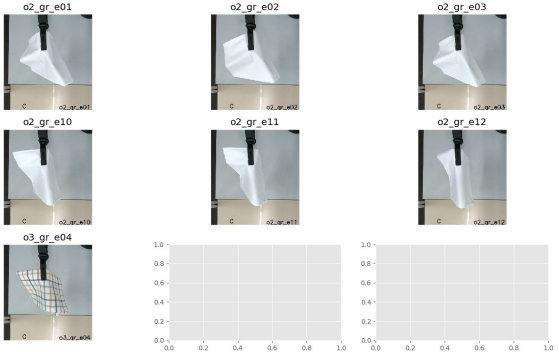


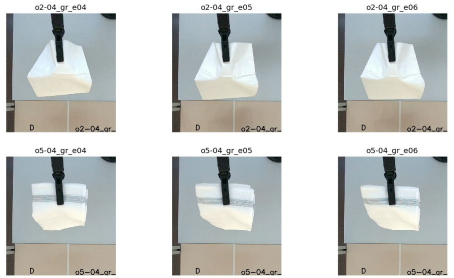


not filling peso 8x8

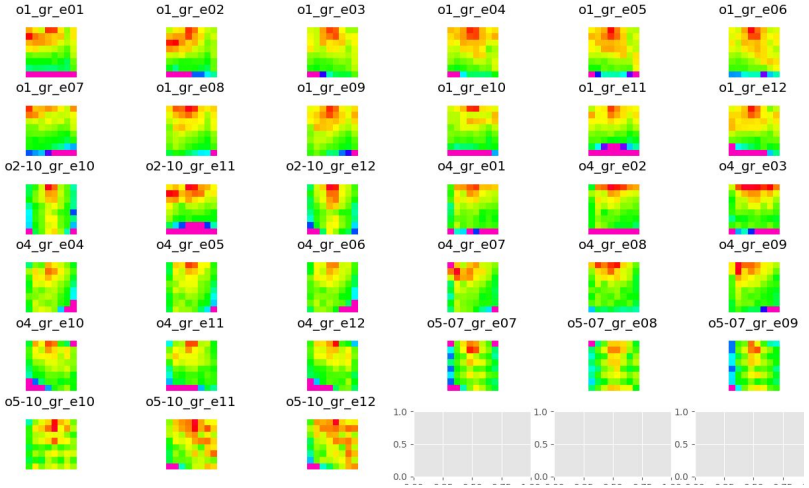


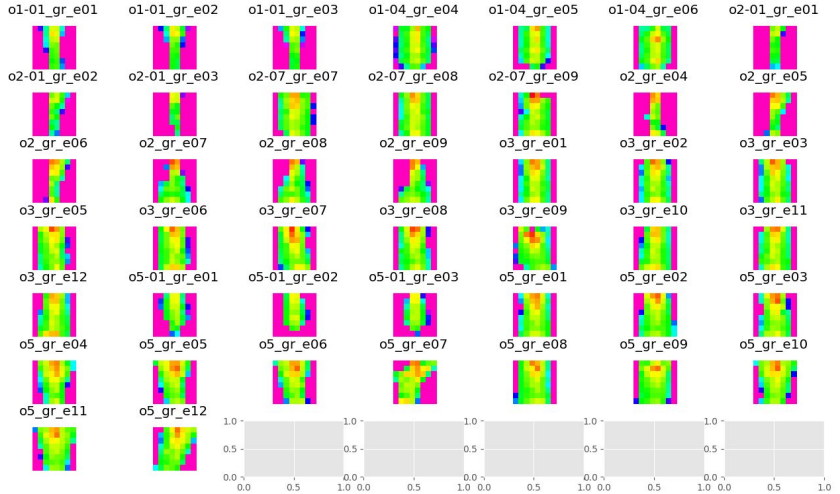


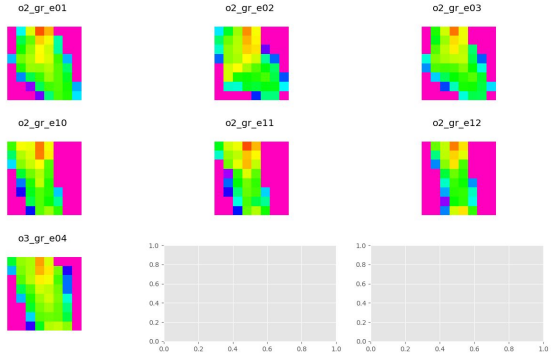


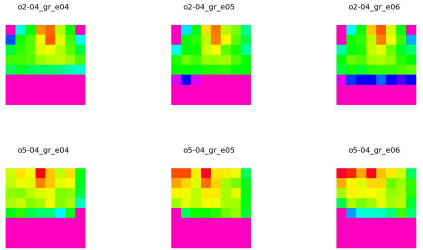


not filling peso 8x8









Changed success extraction method - filling

Combining class label assignment and selecting the one that maximises total success

- n_clusters=4, init=**'random'**, n_init=20, max_iter=500, tol=1e-8, random_state=0

	2x2	3x3	4x4	5x5	6x6	7x7	8x8	9x9	10x10
Success_A	0	100	96.429	100	100	100	100	100	100
Success_B	66.667	0	60	0	55.556	0	26.667	55.556	26.667
Success_C	63.636	72.727	27.273	72.727	0	27.273	72.727	27.273	72.727
Success_D	83.333	0	100	0	0	0	100	100	100
TOTAL	46.667	40	70	40	58.889	34.444	60	68.889	60

- n_clusters=4, init=**k-means++**, n_init=20, max_iter=500, tol=1e-8, random_state=0

	2x2	3x3	4x4	5x5	6x6	7x7	8x8	9x9	10x10
Success_A	100	100	96.429	100	100	100	100	100	100
Success_B	20	0	26.667	60	55.556	57.778	0	55.556	0
Success_C	63.636	72.727	72.727	27.273	27.273	0	27.273	27.273	72.727
Success_D	100	0	100	100	100	0	0	100	0
TOTAL	55.556	40	58.889	71.111	68.889	60	34.444	68.889	40

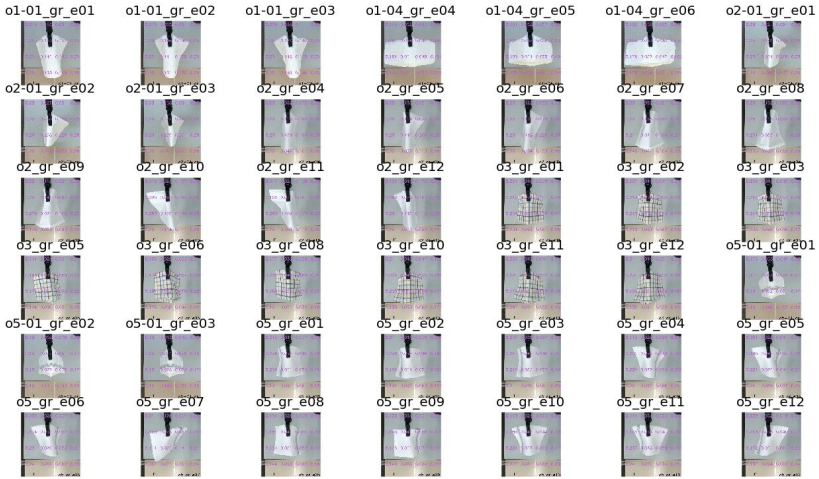
- n_clusters=4, **init=[init_centr]** , n_init=20, max_iter=500, tol=1e-8, **random_state=None** -> samples with LESS intra dist to centroid of class

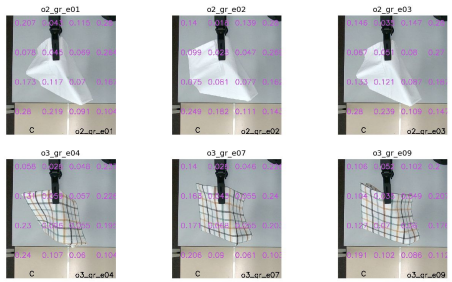
resulting clusters
in next slides

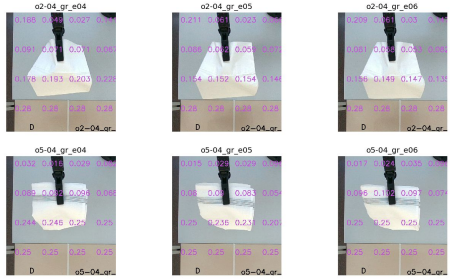
	2x2	3x3	4x4	5x5	6x6	7x7	8x8	9x9	10x10
Success_A	100	100	100	100	100	100	100	100	100
Success_B	28.889	73.333	82.222	60	55.556	55.556	55.556	55.556	55.556
Success_C	54.545	63.636	54.545	27.273	45.455	27.273	27.273	45.455	27.273
Success_D	83.333	100	100	100	100	100	100	100	100
TOTAL	57.778	82.222	85.556	71.111	71.111	68.889	68.889	71.111	68.889

filling 4x4

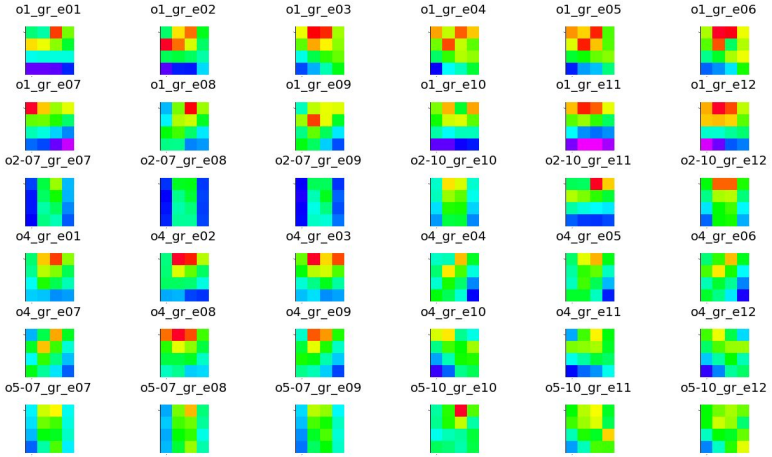

o1_gr_e01
o1_gr_e02
o1_gr_e03
o1_gr_e04
o1_gr_e05
o1_gr_e06
o1_gr_e07
o1_gr_e08
o1_gr_e09
o1_gr_e10
o1_gr_e11
o1_gr_e12
o2-07_gr_e07
o2-07_gr_e08
o2-07_gr_e09
o2-10_gr_e10
o2-10_gr_e11
o2-10_gr_e12
o4_gr_e01
o4_gr_e02
o4_gr_e03
o4_gr_e04
o4_gr_e05
o4_gr_e06
o4_gr_e07
o4_gr_e08
o4_gr_e09
o4_gr_e10
o4_gr_e11
o4_gr_e12
o5-07_gr_e07
o5-07_gr_e08
o5-07_gr_e09
o5-10_gr_e10
o5-10_gr_e11
o5-10_gr_e12

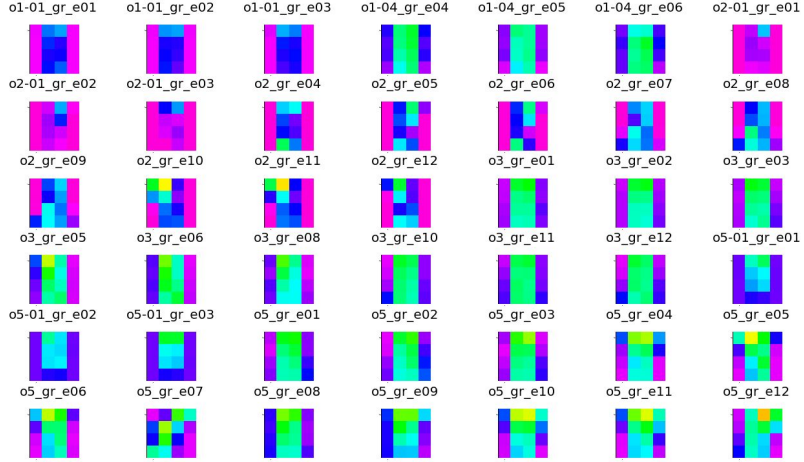

o1-01_gr_e01
o1-01_gr_e02
o1-01_gr_e03
o1-04_gr_e04
o1-04_gr_e05
o1-04_gr_e06
o2-01_gr_e01
o2-01_gr_e02
o2-01_gr_e03
o2_gr_e04
o2_gr_e05
o2_gr_e06
o2_gr_e07
o2_gr_e08
o2_gr_e09
o2_gr_e10
o2_gr_e11
o2_gr_e12
o3_gr_e01
o3_gr_e02
o3_gr_e03
o3_gr_e05
o3_gr_e06
o3_gr_e08
o3_gr_e10
o3_gr_e11
o3_gr_e12
o5-01_gr_e01
o5-01_gr_e02
o5-01_gr_e03
o5_gr_e01
o5_gr_e02
o5_gr_e03
o5_gr_e04
o5_gr_e05
o5_gr_e06
o5_gr_e07
o5_gr_e08
o5_gr_e09
o5_gr_e10
o5_gr_e11
o5_gr_e12

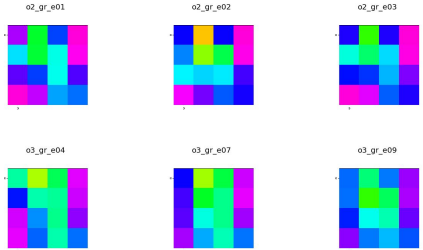

o2_gr_e01
o2_gr_e02
o2_gr_e03
o3_gr_e04
o3_gr_e07
o3_gr_e09

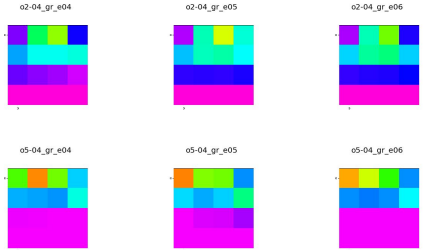

o2-04_gr_e04
o2-04_gr_e05
o2-04_gr_e06
o5-04_gr_e04
o5-04_gr_e05
o5-04_gr_e06

filling 4x4


o1_gr_e01
o1_gr_e02
o1_gr_e03
o1_gr_e04
o1_gr_e05
o1_gr_e06
o1_gr_e07
o1_gr_e08
o1_gr_e09
o1_gr_e10
o1_gr_e11
o1_gr_e12
o2-07_gr_e07
o2-07_gr_e08
o2-07_gr_e09
o2-10_gr_e10
o2-10_gr_e11
o2-10_gr_e12
o4_gr_e01
o4_gr_e02
o4_gr_e03
o4_gr_e04
o4_gr_e05
o4_gr_e06
o4_gr_e07
o4_gr_e08
o4_gr_e09
o4_gr_e10
o4_gr_e11
o4_gr_e12
o5-07_gr_e07
o5-07_gr_e08
o5-07_gr_e09
o5-10_gr_e10
o5-10_gr_e11
o5-10_gr_e12


o1-01_gr_e01
o1-01_gr_e02
o1-01_gr_e03
o1-04_gr_e04
o1-04_gr_e05
o1-04_gr_e06
o2-01_gr_e01
o2-01_gr_e02
o2-01_gr_e03
o2_gr_e04
o2_gr_e05
o2_gr_e06
o2_gr_e07
o2_gr_e08
o2_gr_e09
o2_gr_e10
o2_gr_e11
o2_gr_e12
o3_gr_e01
o3_gr_e02
o3_gr_e03
o3_gr_e05
o3_gr_e06
o3_gr_e08
o3_gr_e10
o3_gr_e11
o3_gr_e12
o5-01_gr_e01
o5-01_gr_e02
o5-01_gr_e03
o5_gr_e01
o5_gr_e02
o5_gr_e03
o5_gr_e04
o5_gr_e05
o5_gr_e06
o5_gr_e07
o5_gr_e08
o5_gr_e09
o5_gr_e10
o5_gr_e11
o5_gr_e12


o2_gr_e01
o2_gr_e02
o2_gr_e03
o3_gr_e04
o3_gr_e07
o3_gr_e09


o2-04_gr_e04
o2-04_gr_e05
o2-04_gr_e06
o5-04_gr_e04
o5-04_gr_e05
o5-04_gr_e06

TO DO

- Get more samples of C and D (igualar a A y B)
- Test other clustering
- Modify GT based on intra/inter distances
- Extract error clusterings (for each predicted cluster, compute numbers of samples misclassified)
- Paper Arnau