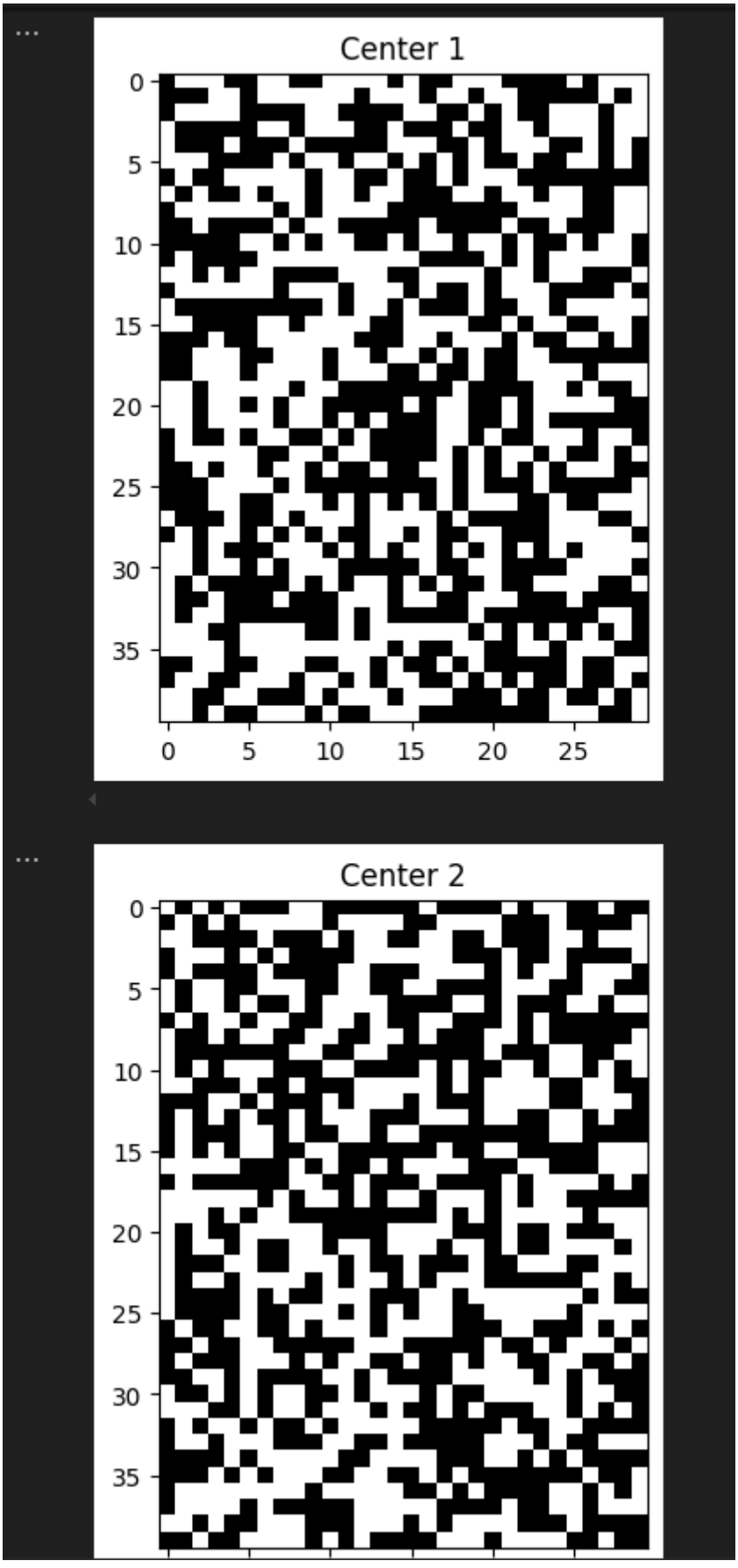
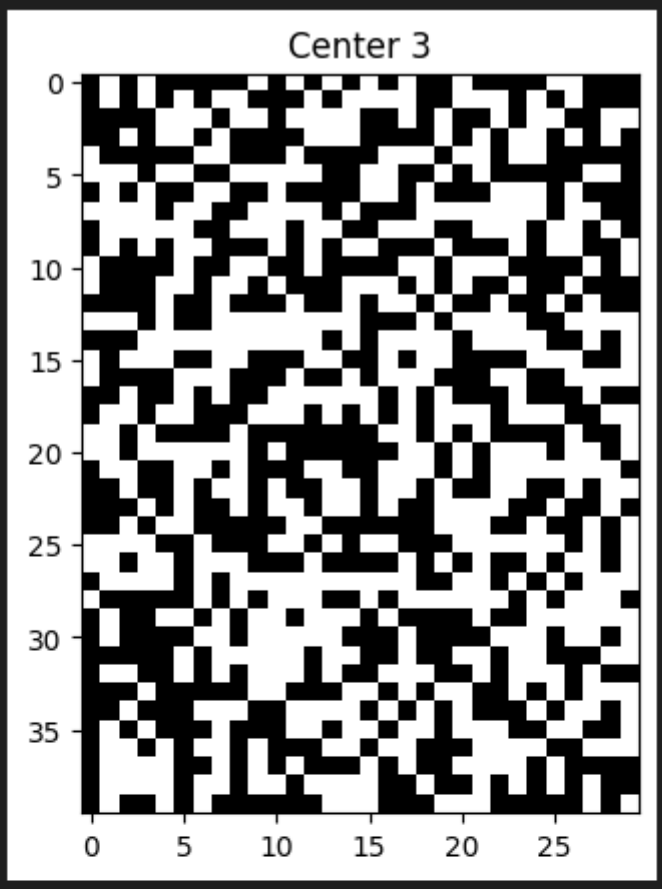
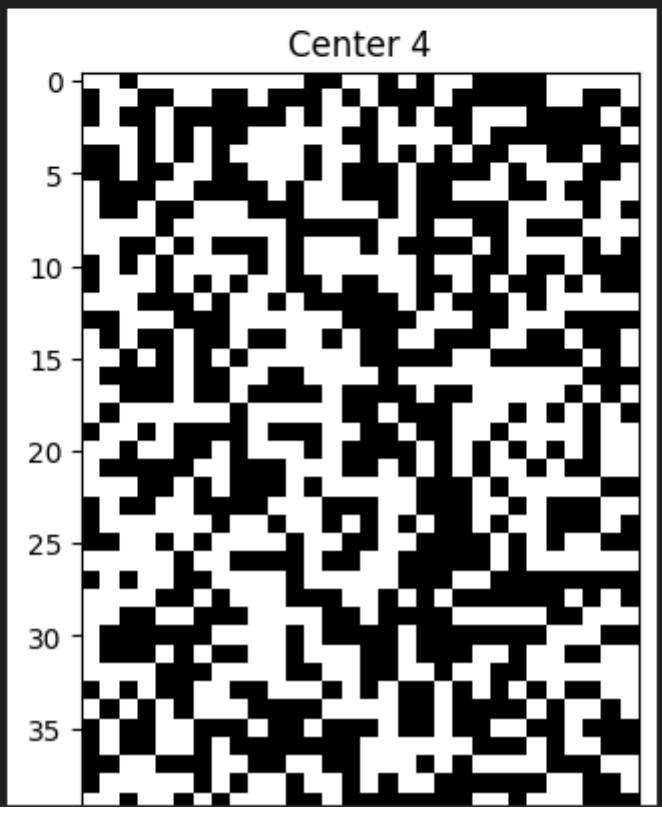
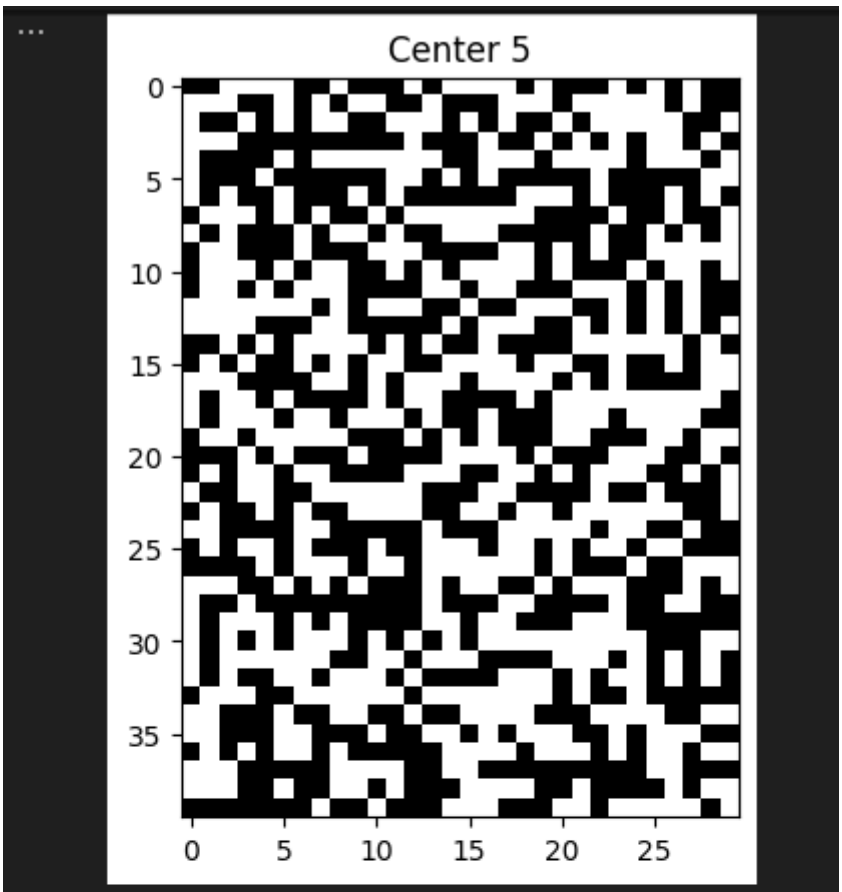
a.I:

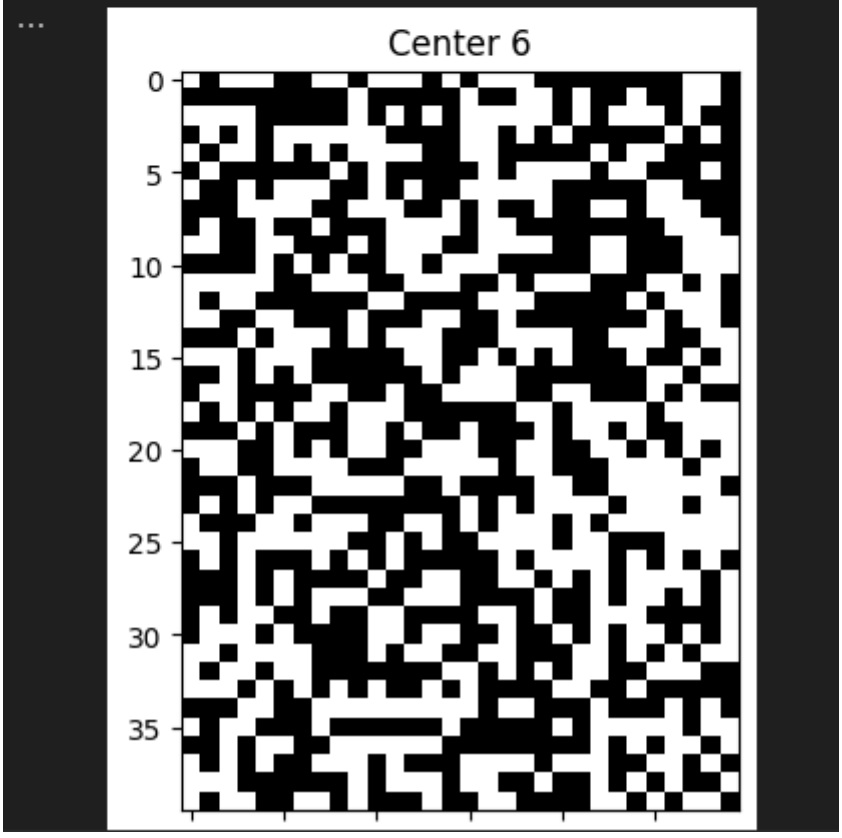
شاخص های تصادفی اولیه

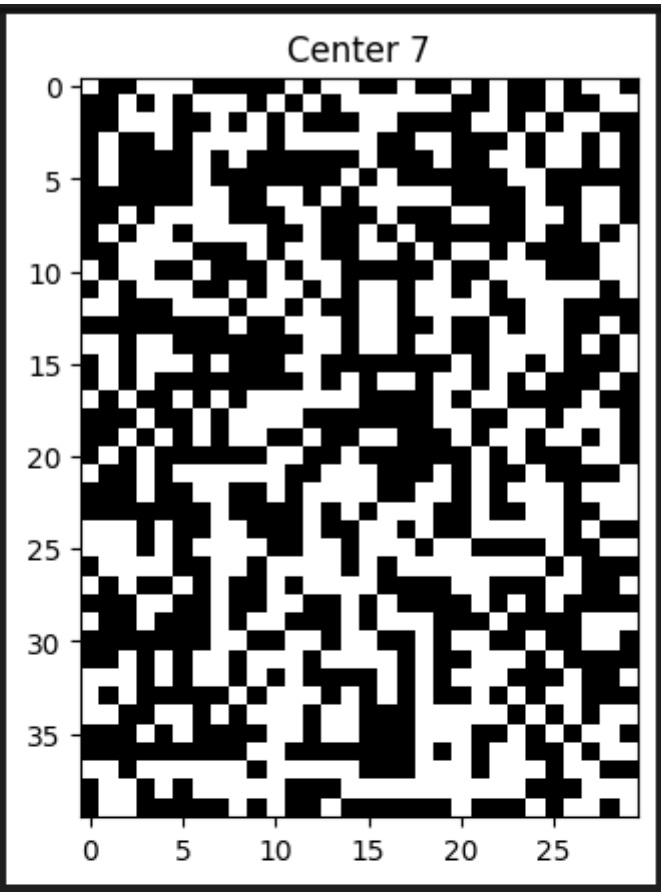


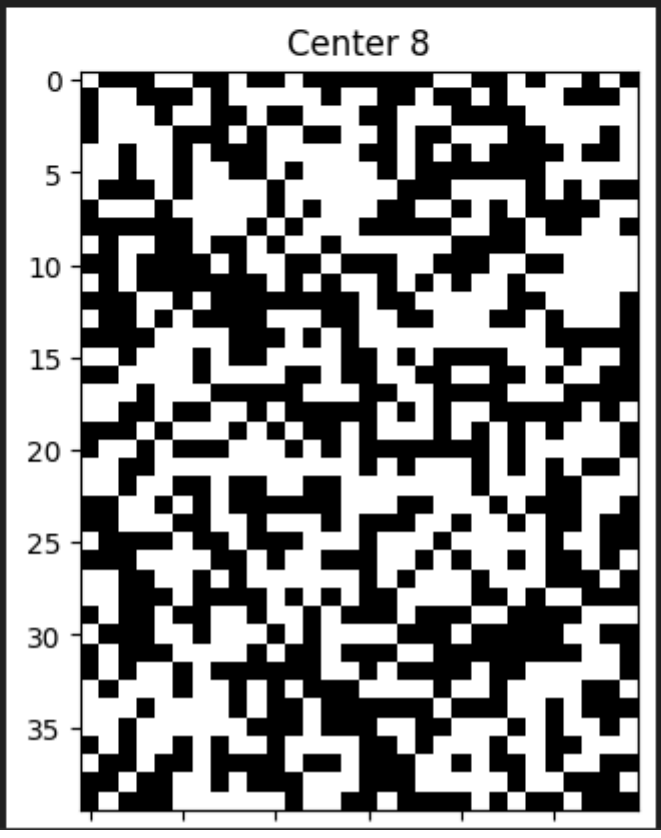

Center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25

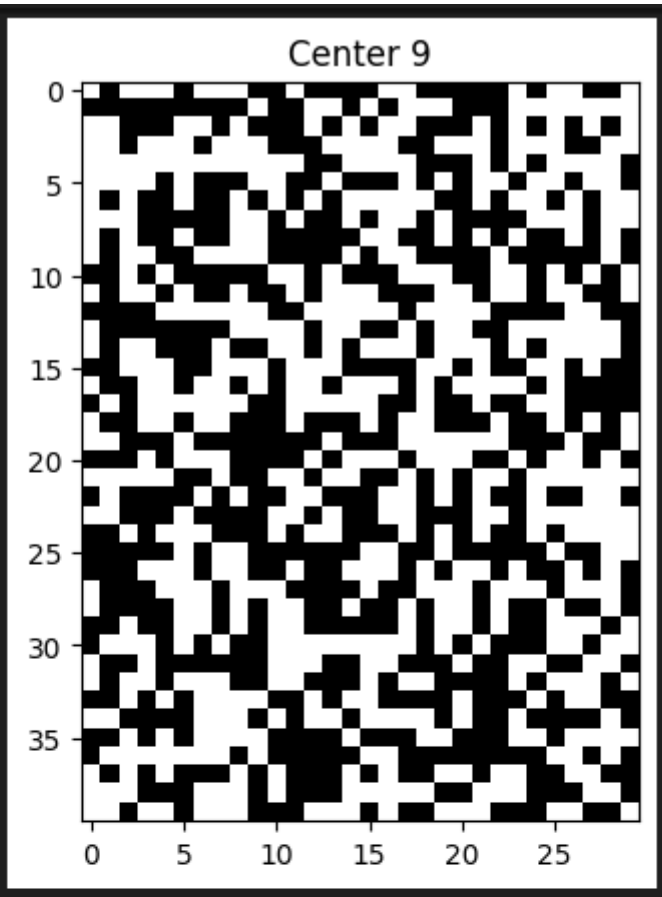

Center 4
0
5
10
15
20
25
30
35

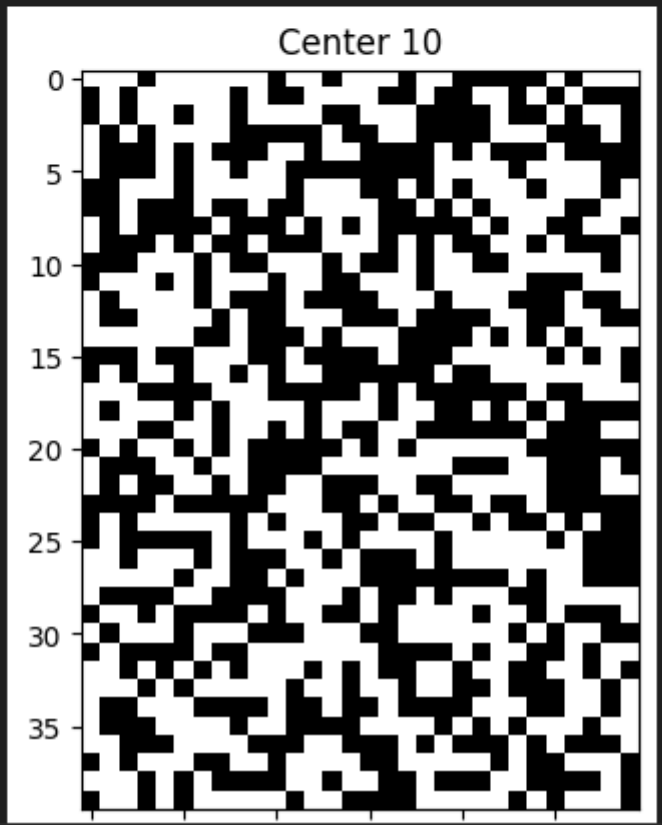

Center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Center 6
0
5
10
15
20
25
30
35

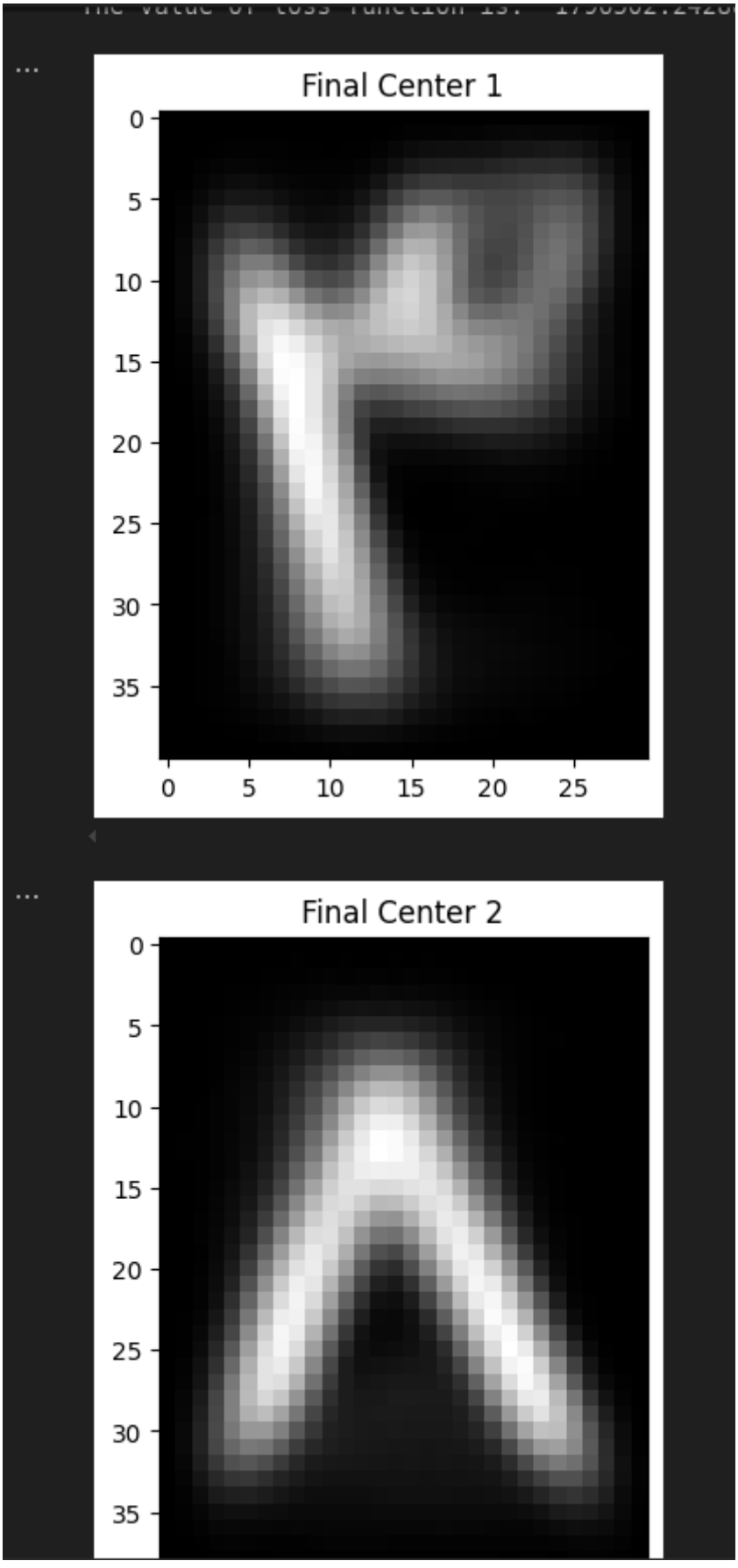

Center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25

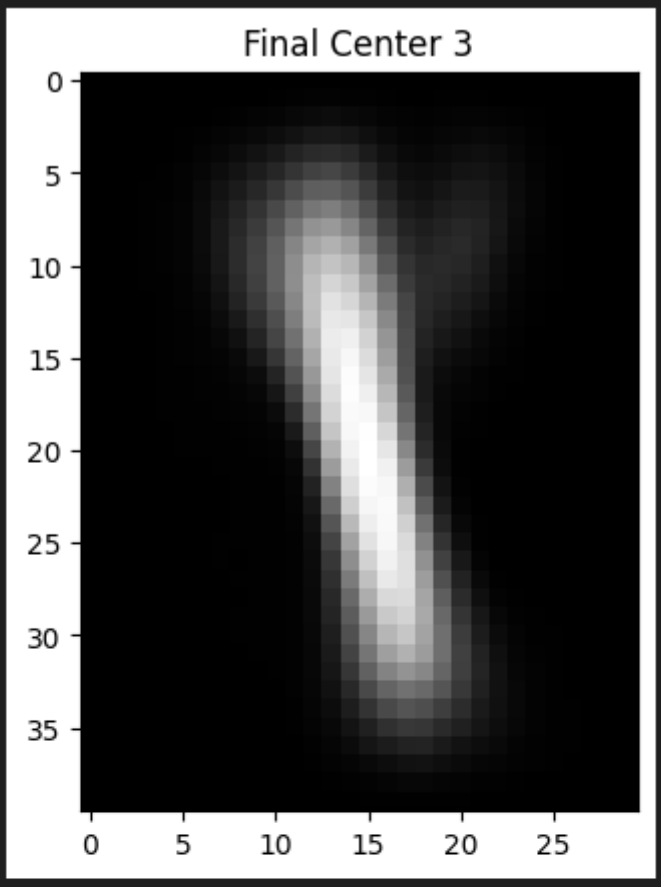

Center 8
0
5
10
15
20
25
30
35

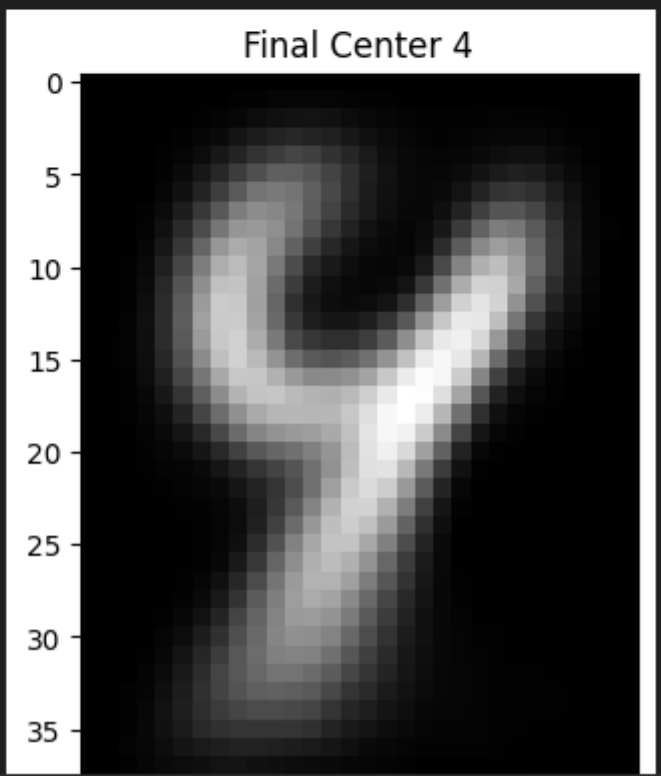

Center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25

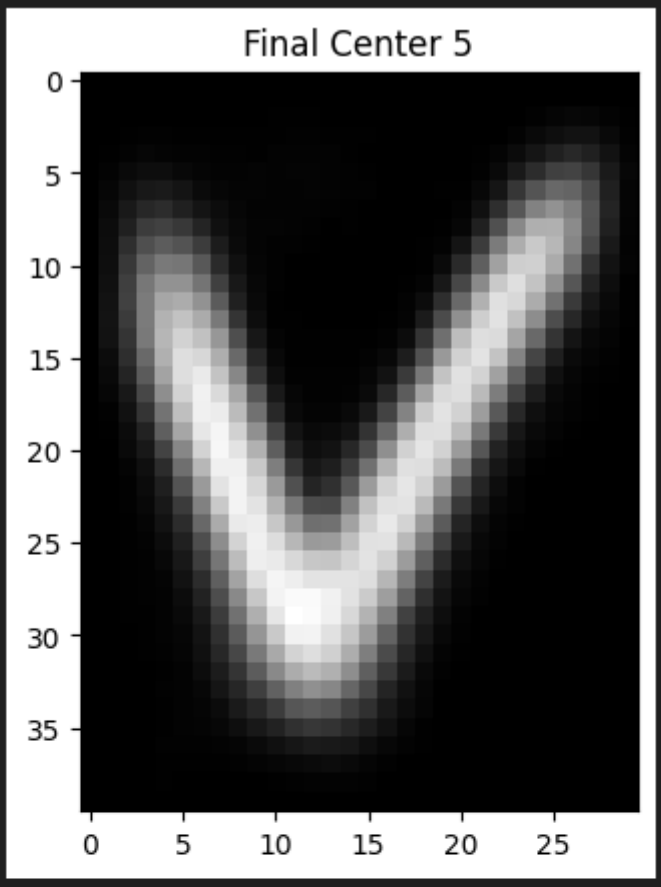

Center 10
0
5
10
15
20
25
30
35

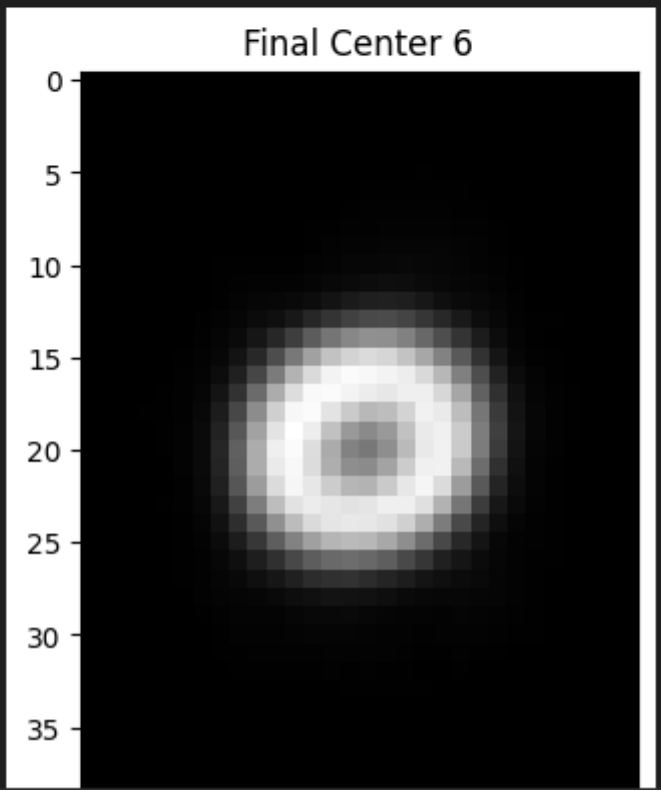
شاخص ها نهایی

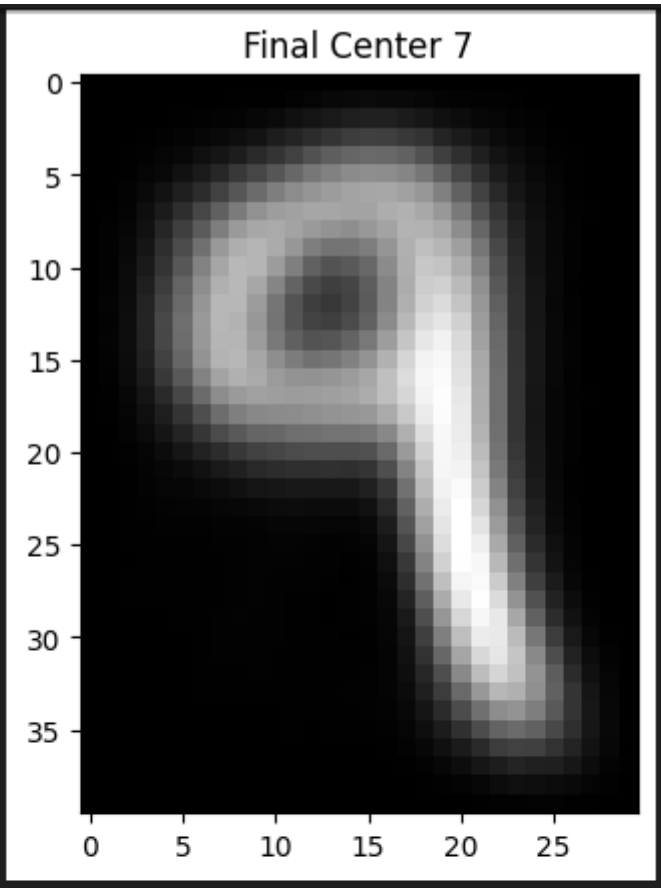


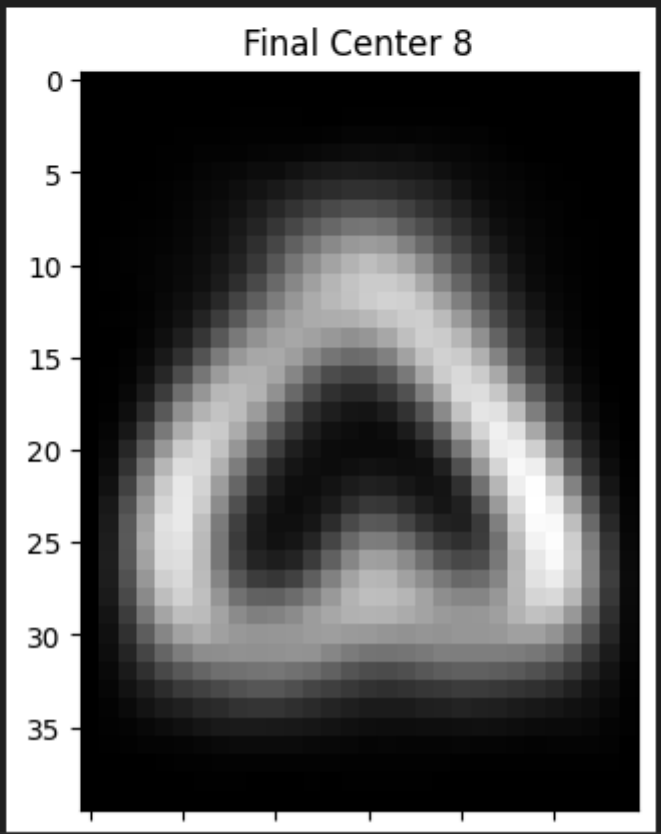

Final Center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25

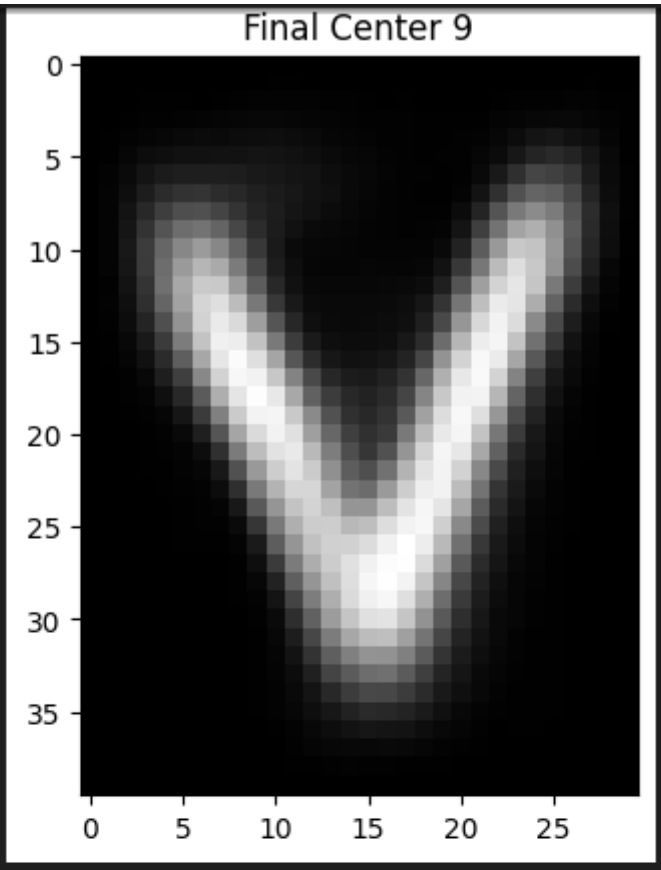

Final Center 4
0
5
10
15
20
25
30
35

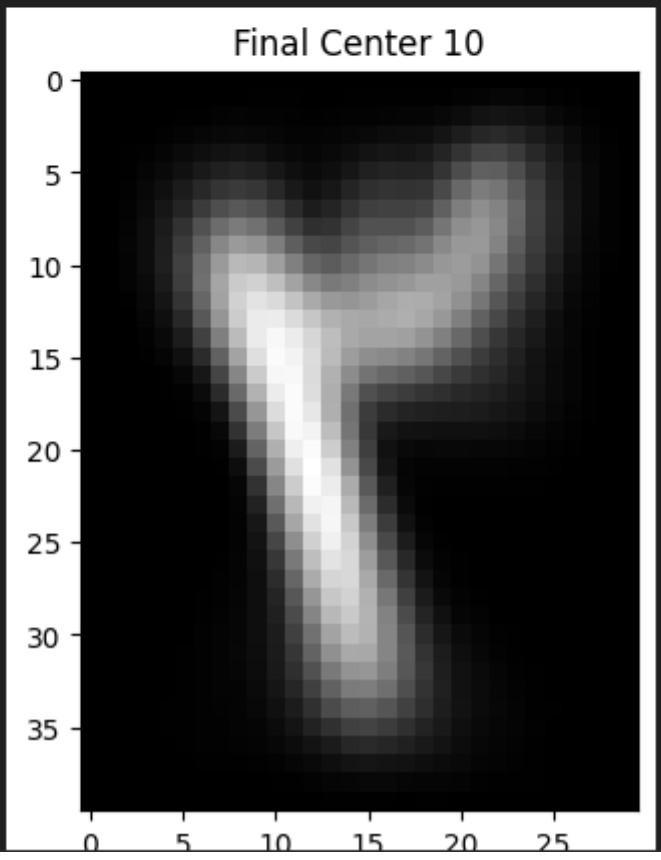

Final Center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 6
0
5
10
15
20
25
30
35

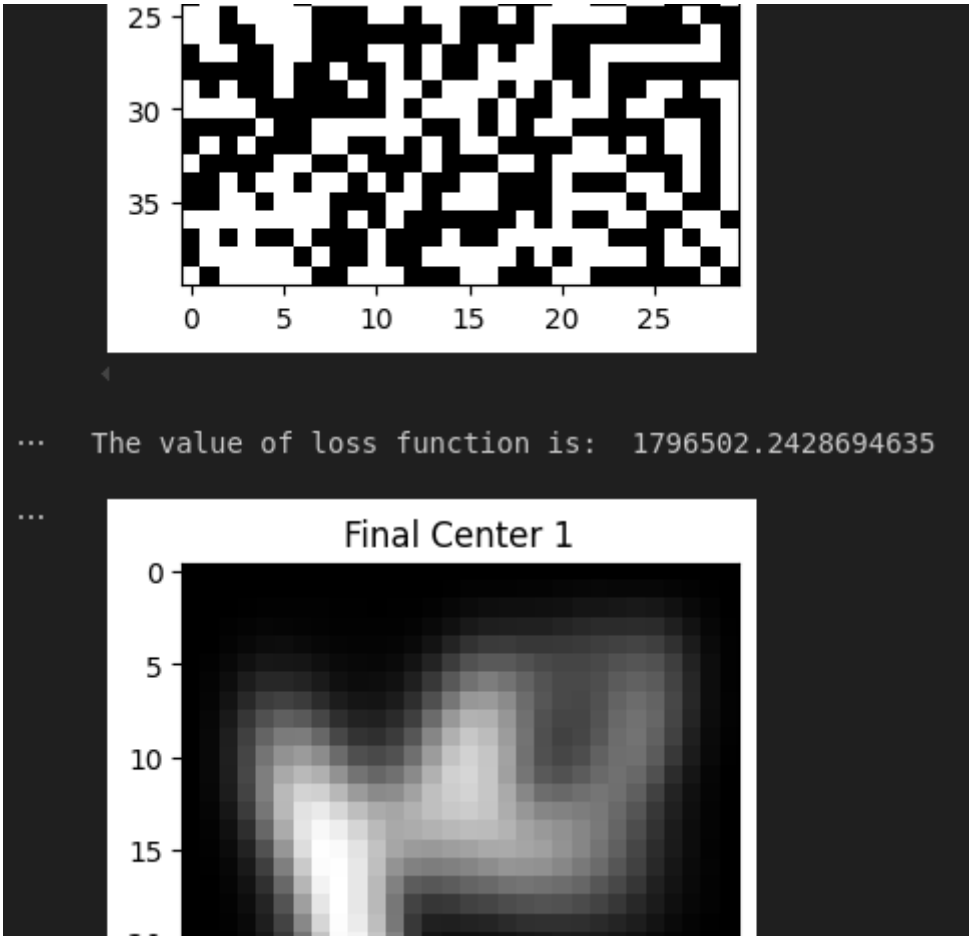

Final Center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 8
0
5
10
15
20
25
30
35


Final Center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25

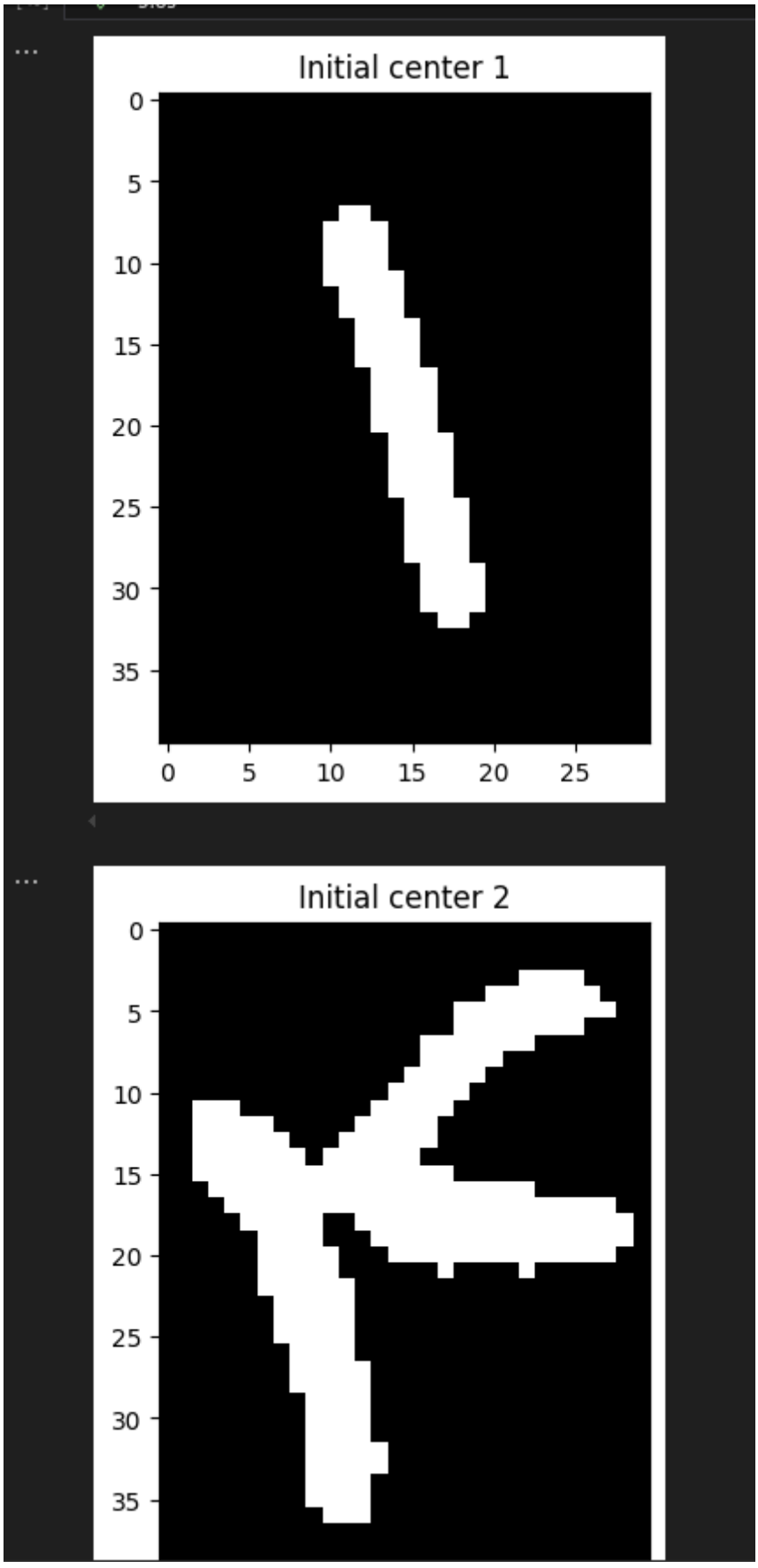

Final Center 10
0
5
10
15
20
25
30
35
0
5
10
15
20
25

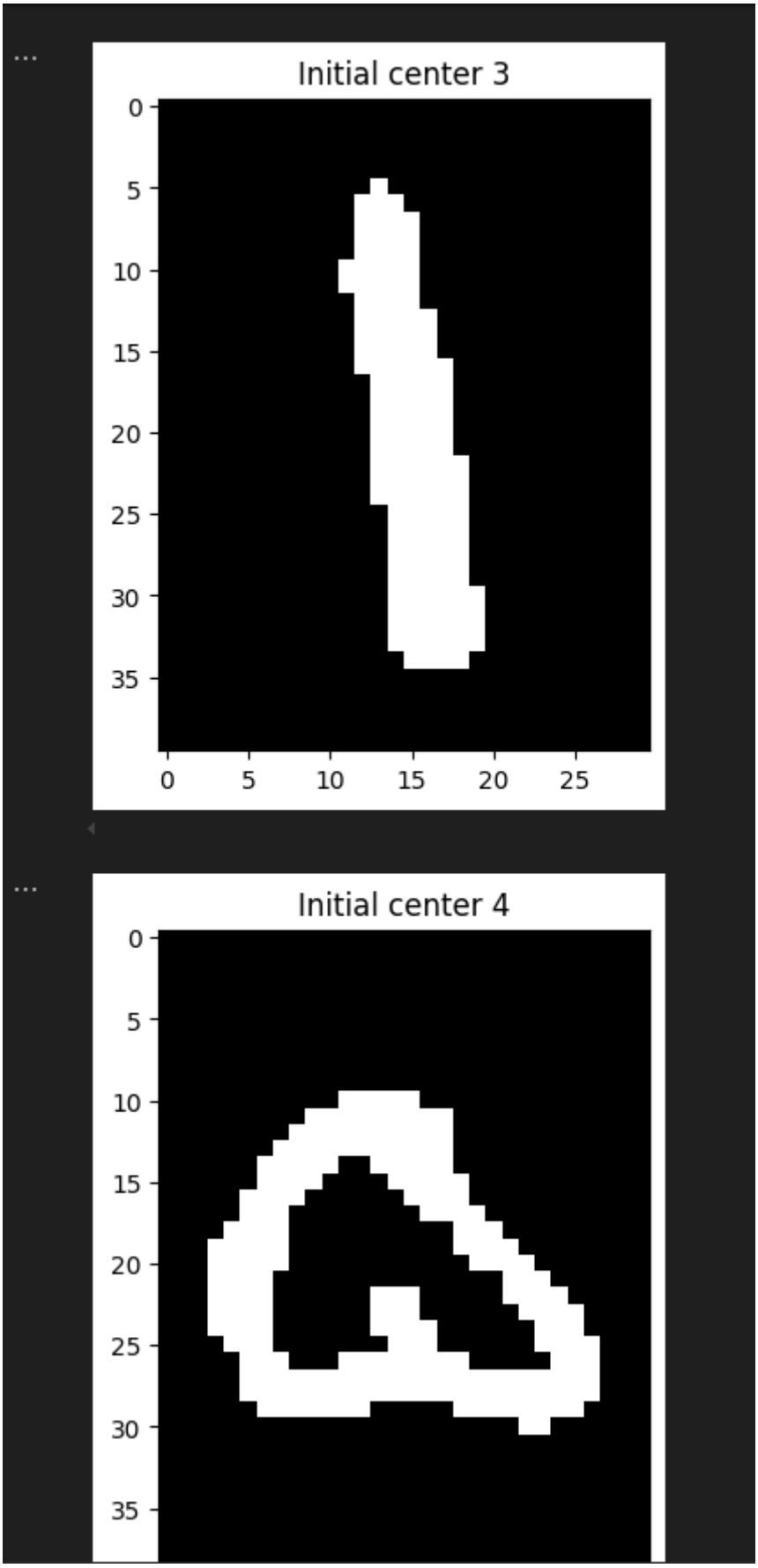
مقدار تابع هزینه

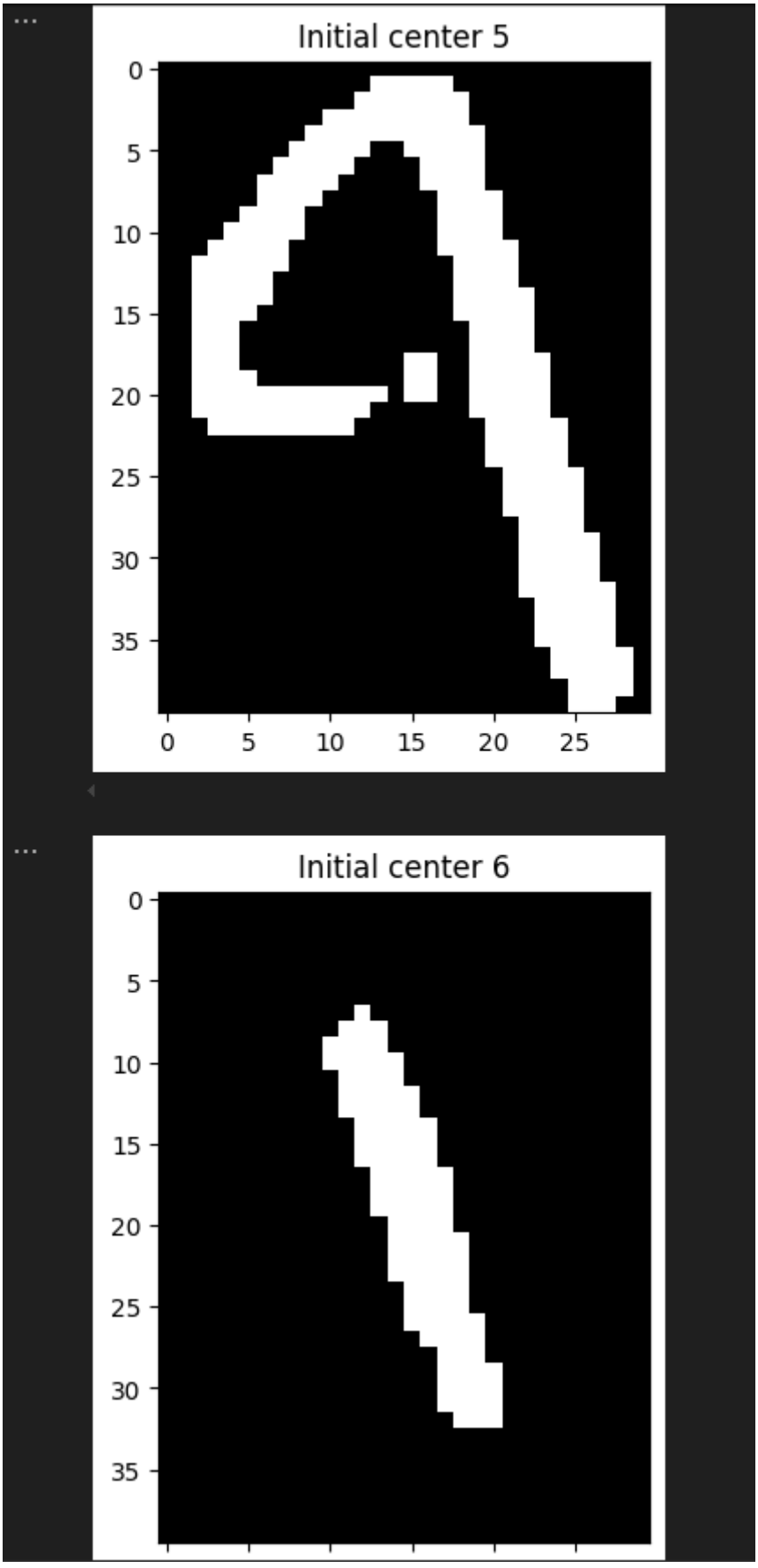


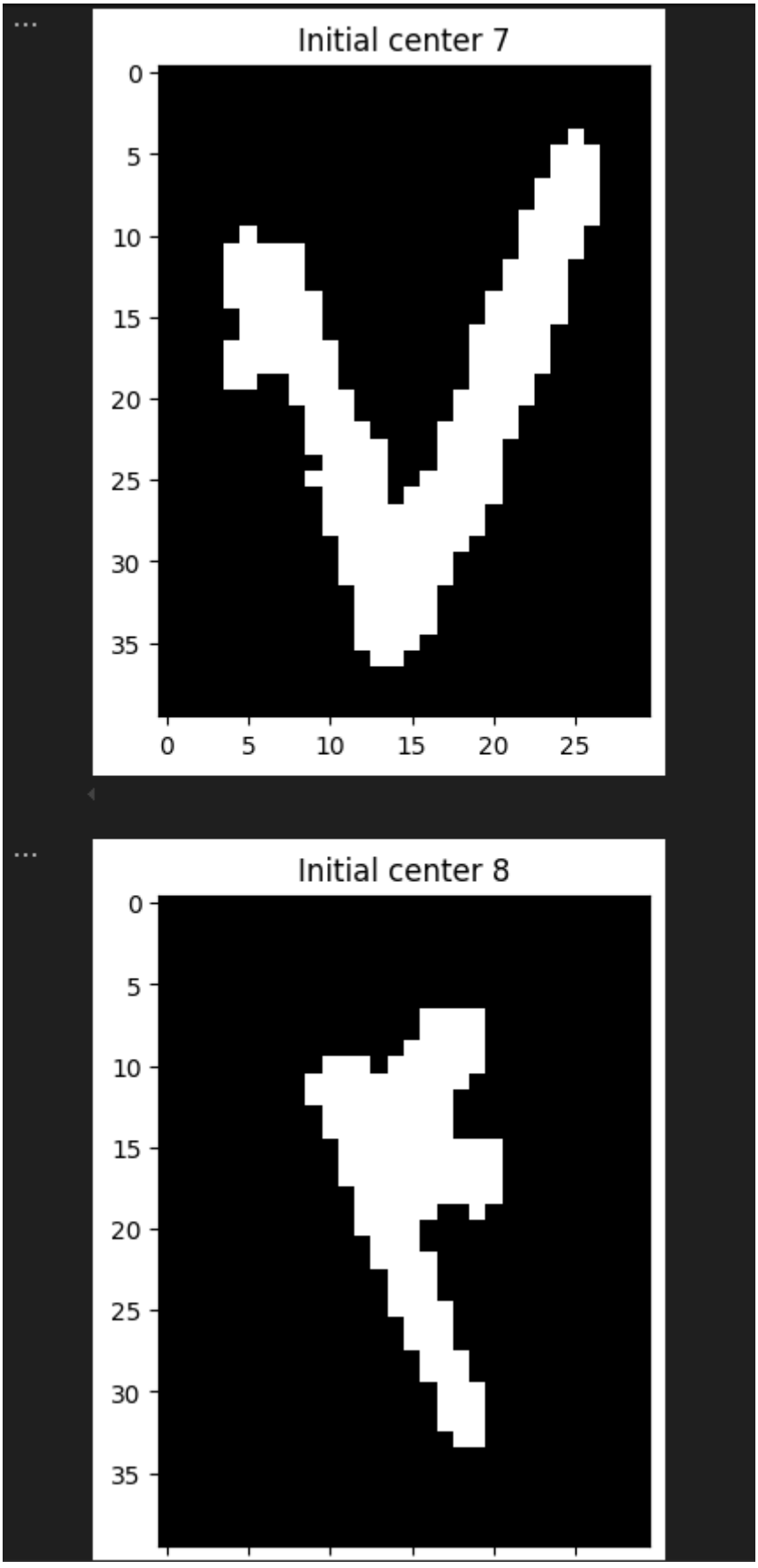
a.II:

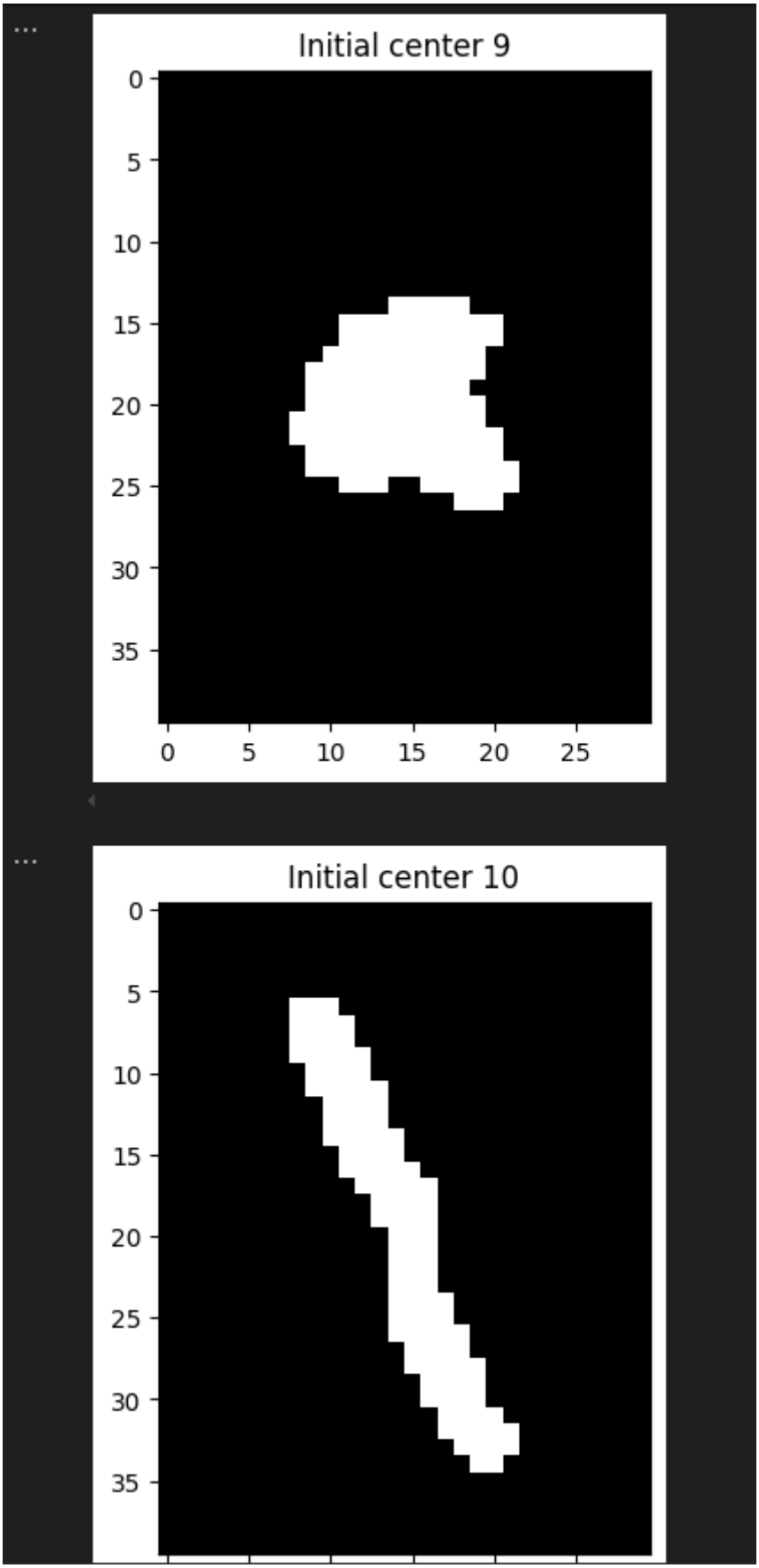
شاخص های تصادفی انتخاب شده از بین داده ها



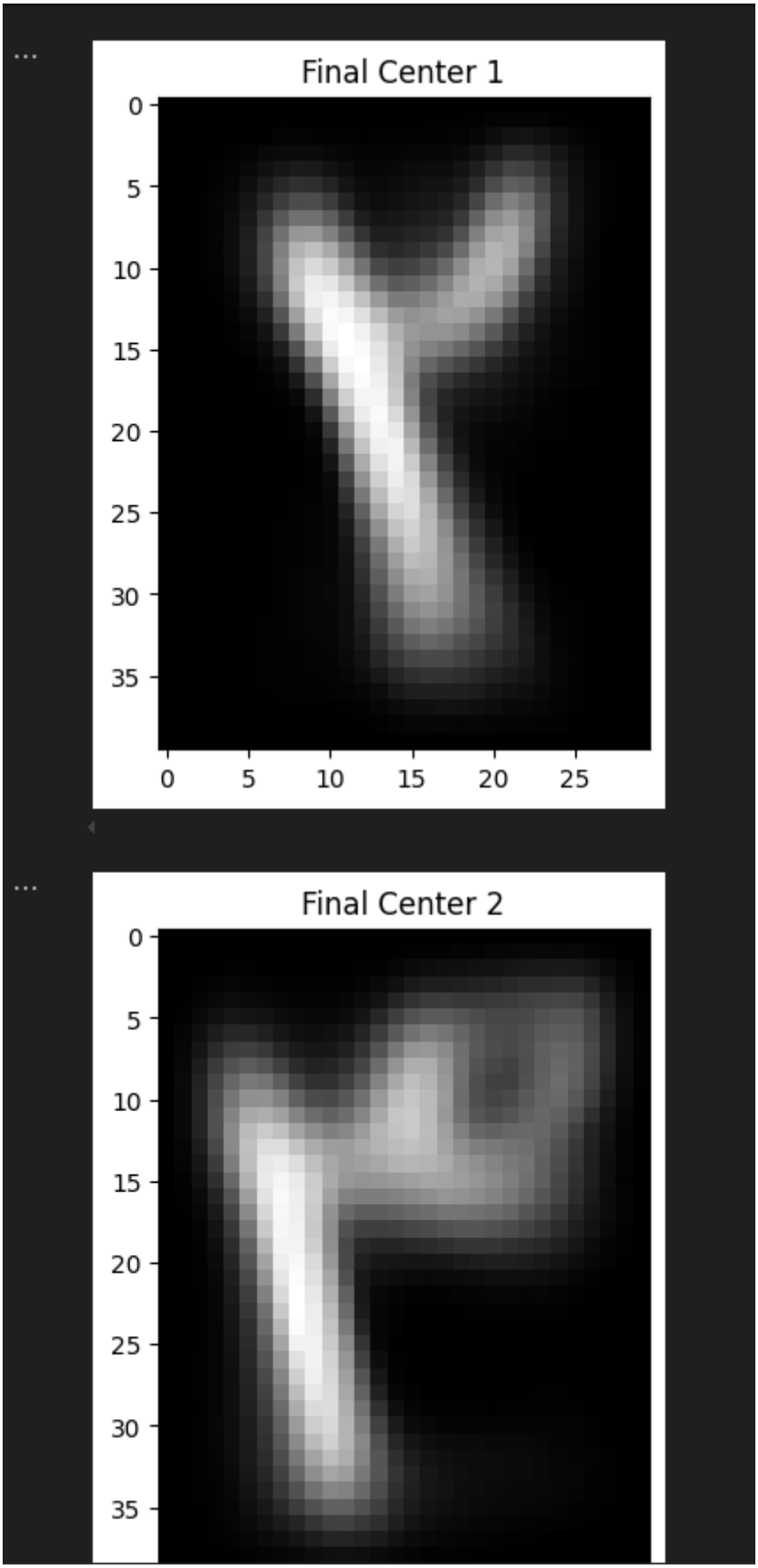

Initial center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial center 4
0
5
10
15
20
25
30
35

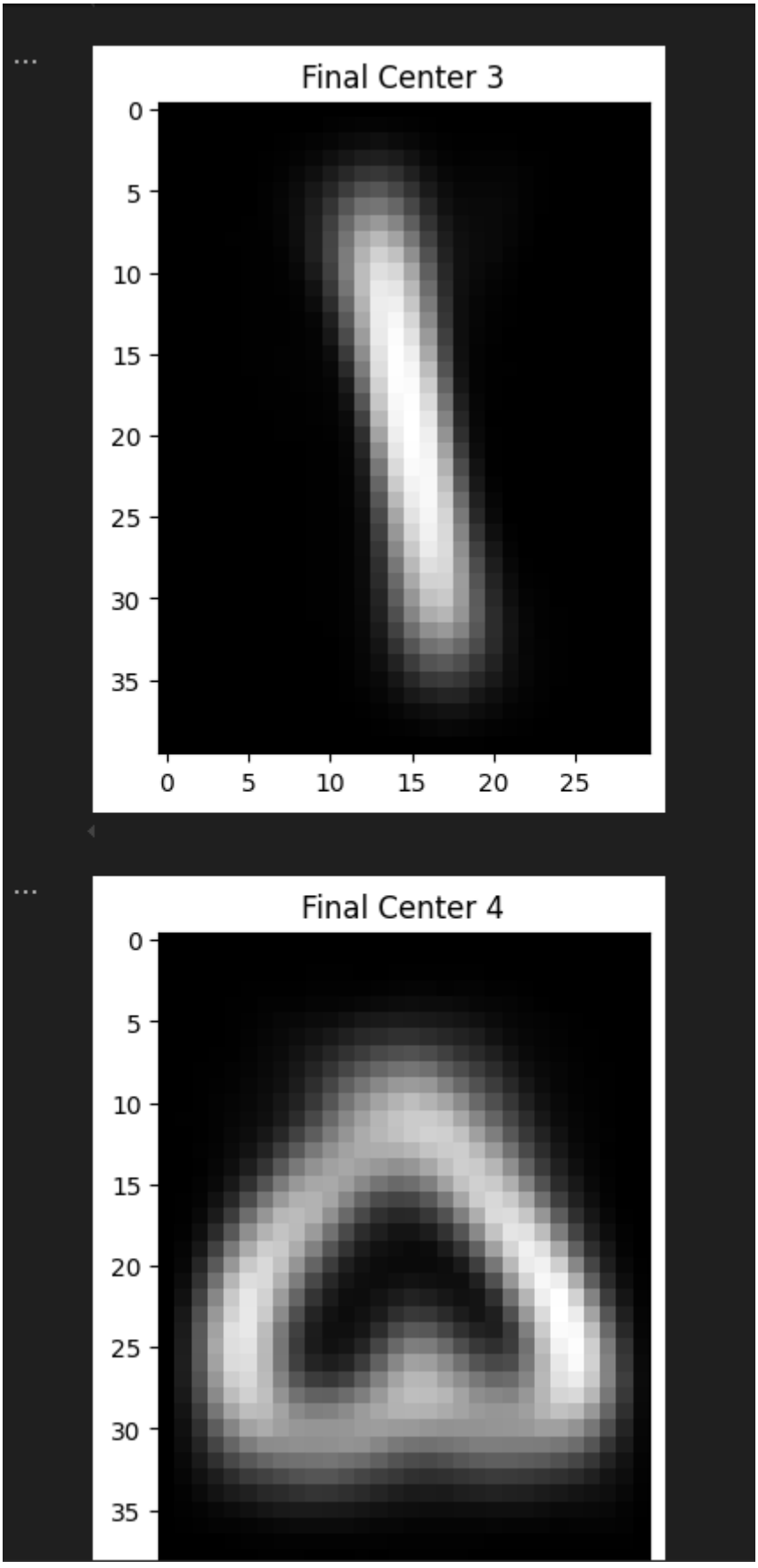

Initial center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial center 6
0
5
10
15
20
25
30
35

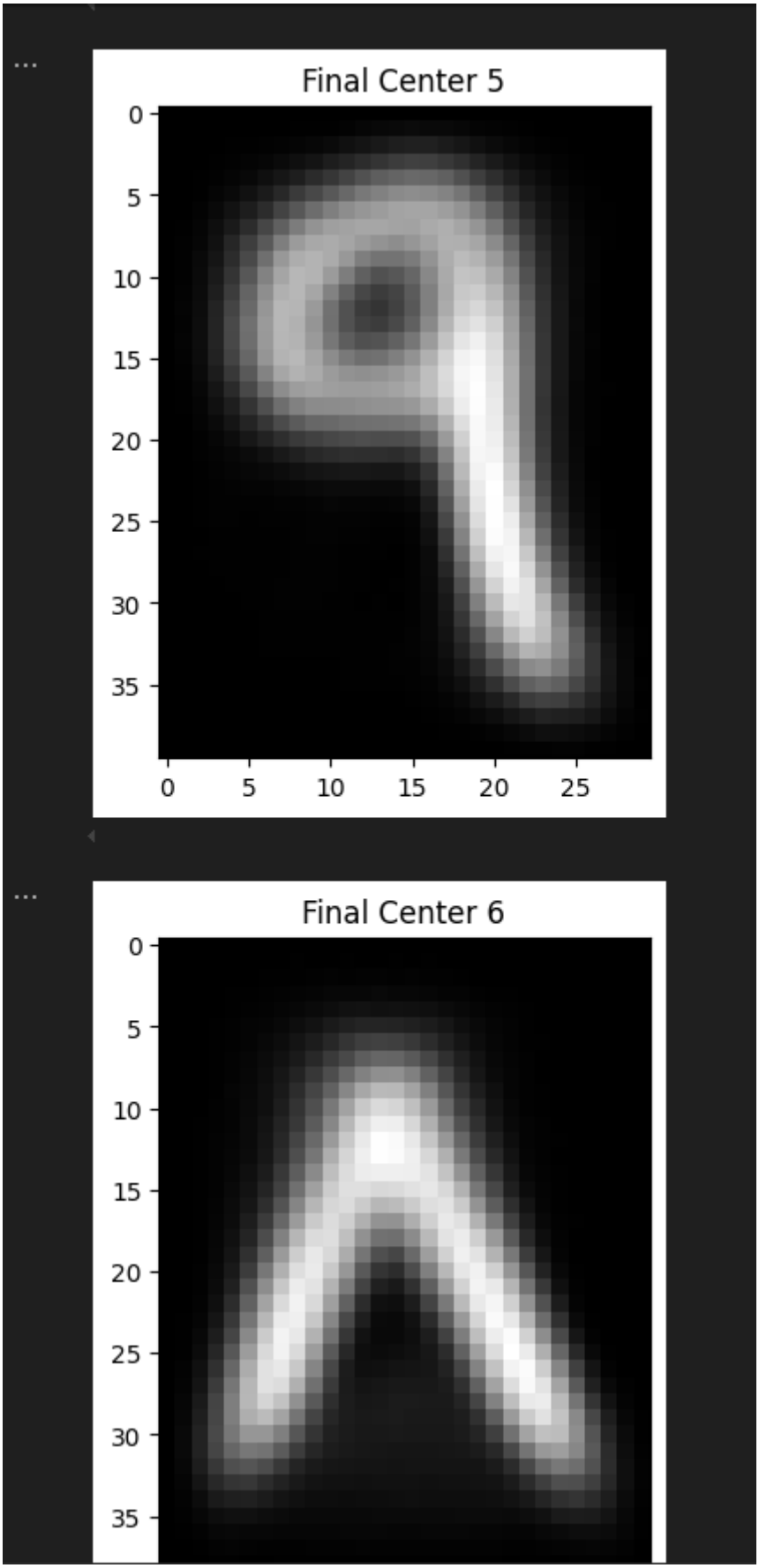

Initial center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial center 8
0
5
10
15
20
25
30
35

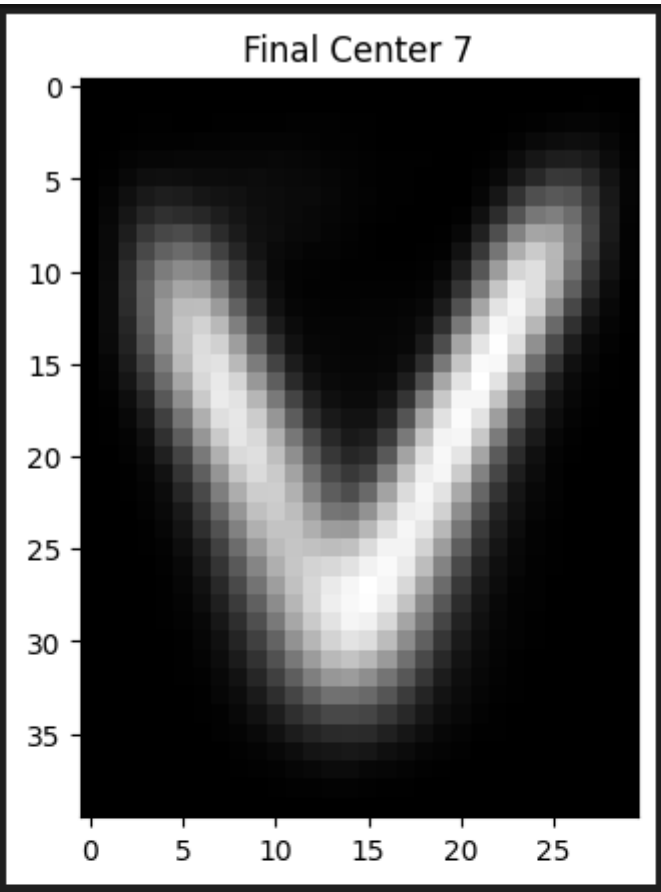

Initial center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial center 10
0
5
10
15
20
25
30
35

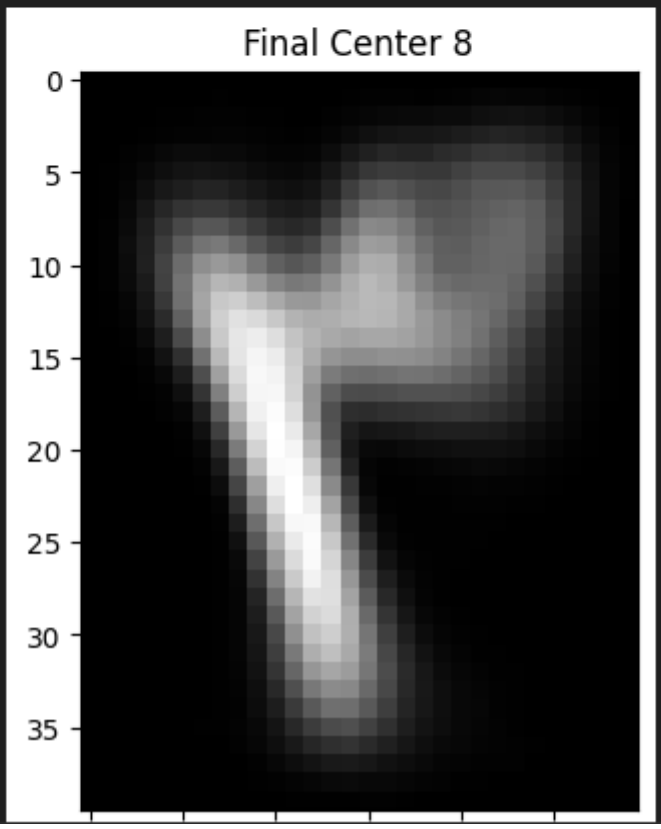
شاخص های نهایی

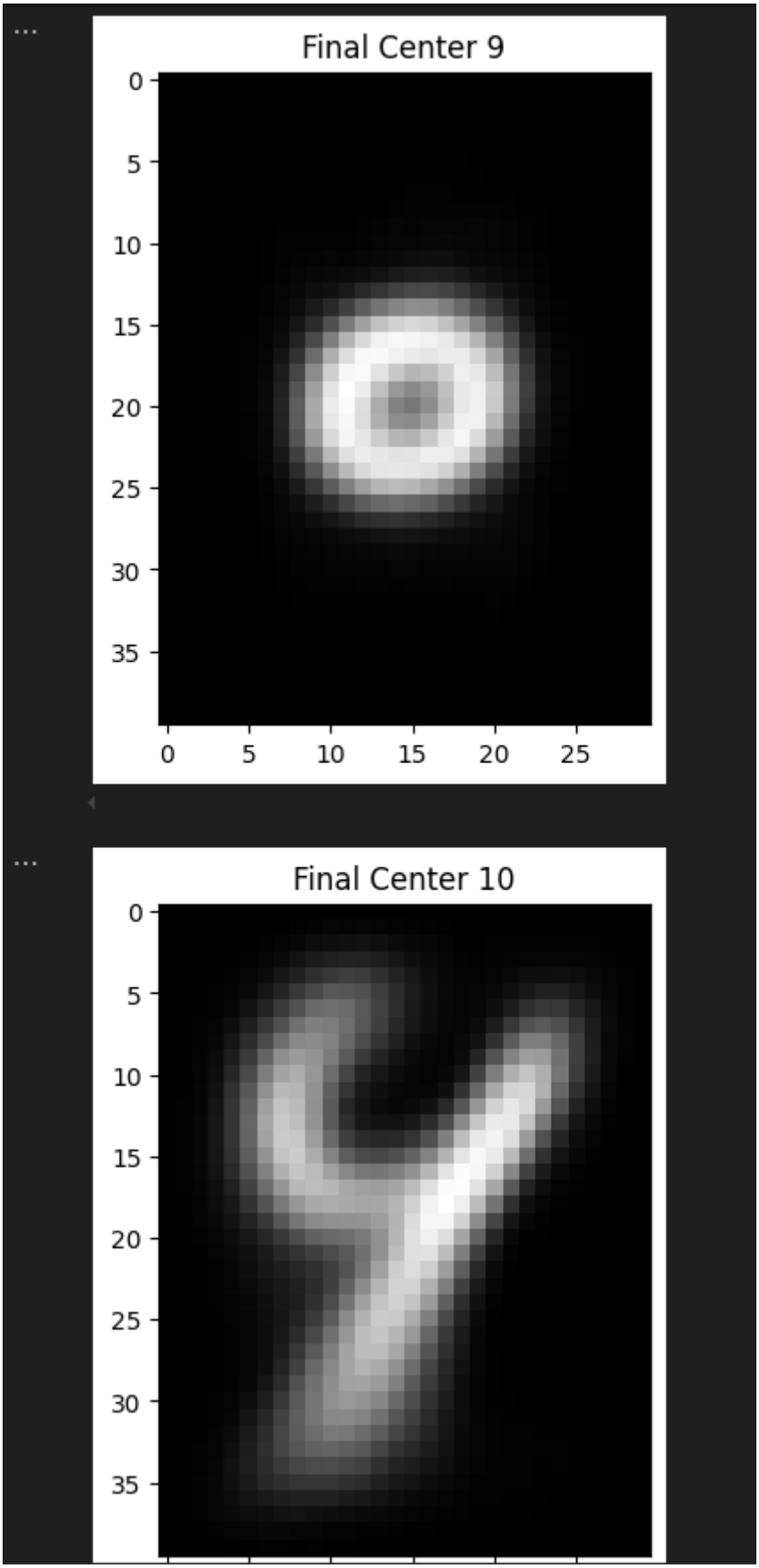



Final Center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 4
0
5
10
15
20
25
30
35

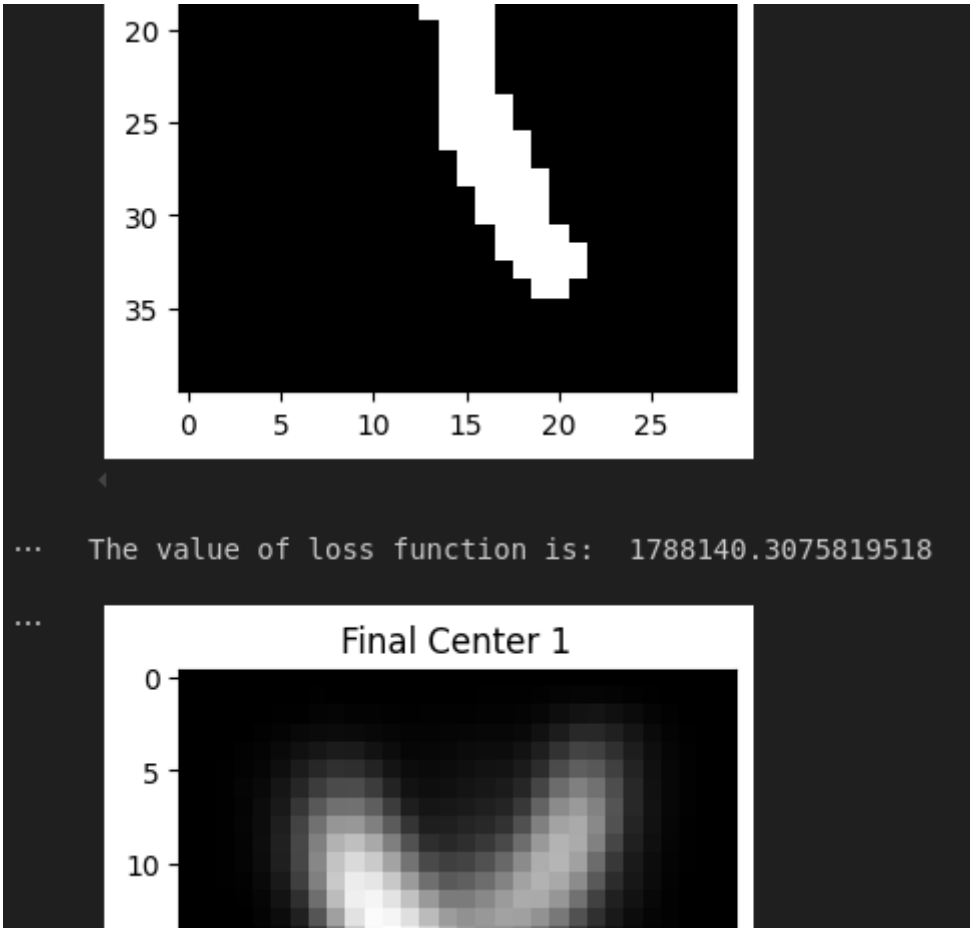

Final Center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 6
0
5
10
15
20
25
30
35


Final Center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 8
0
5
10
15
20
25
30
35

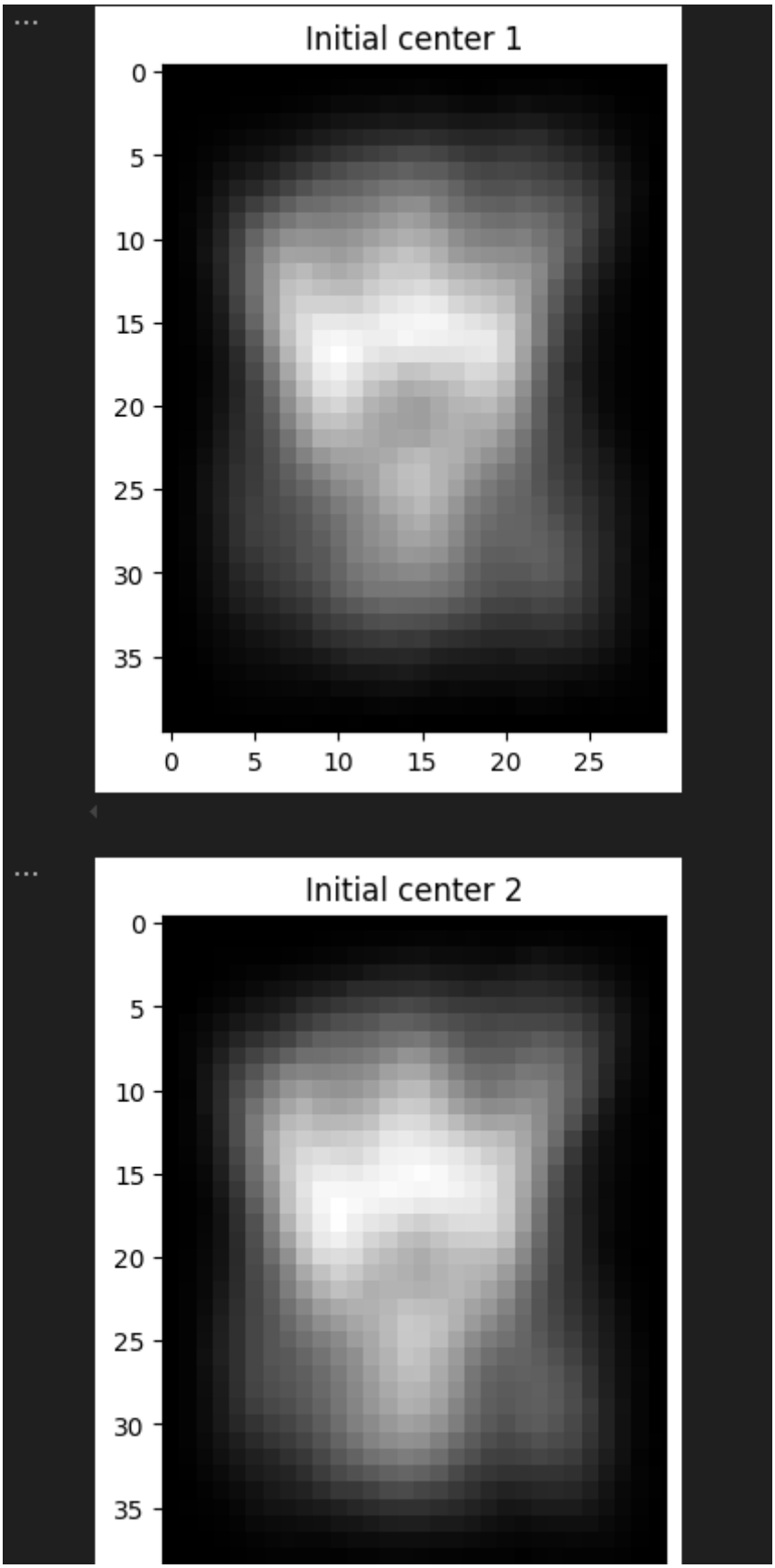

Final Center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 10
0
5
10
15
20
25
30
35

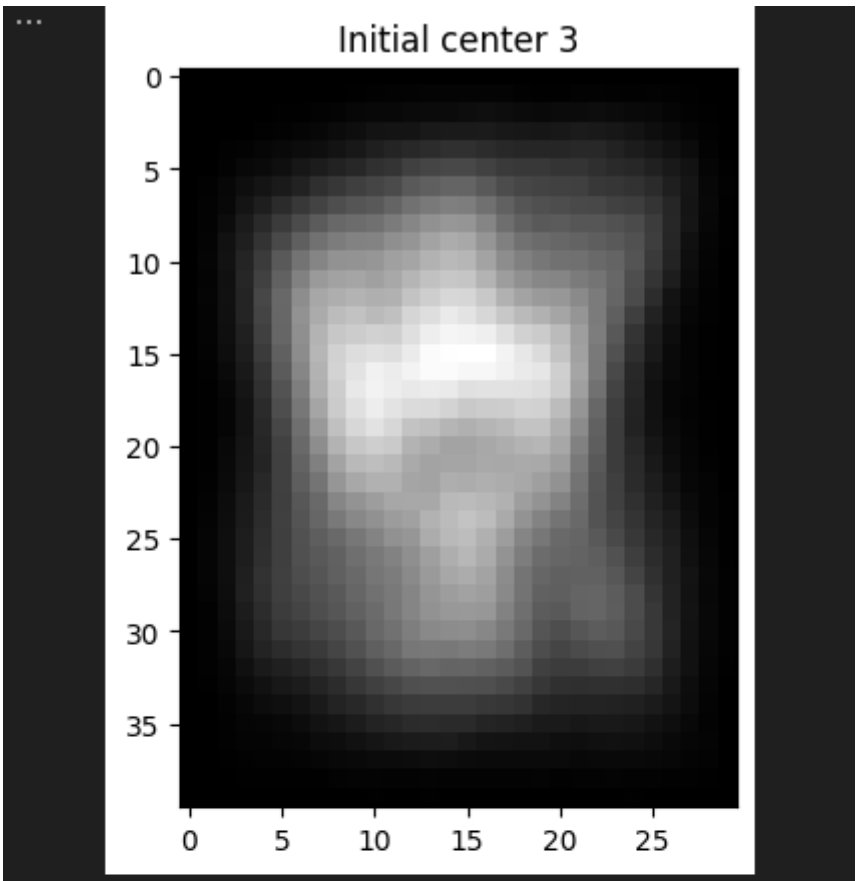
مقدار تابع هزینه

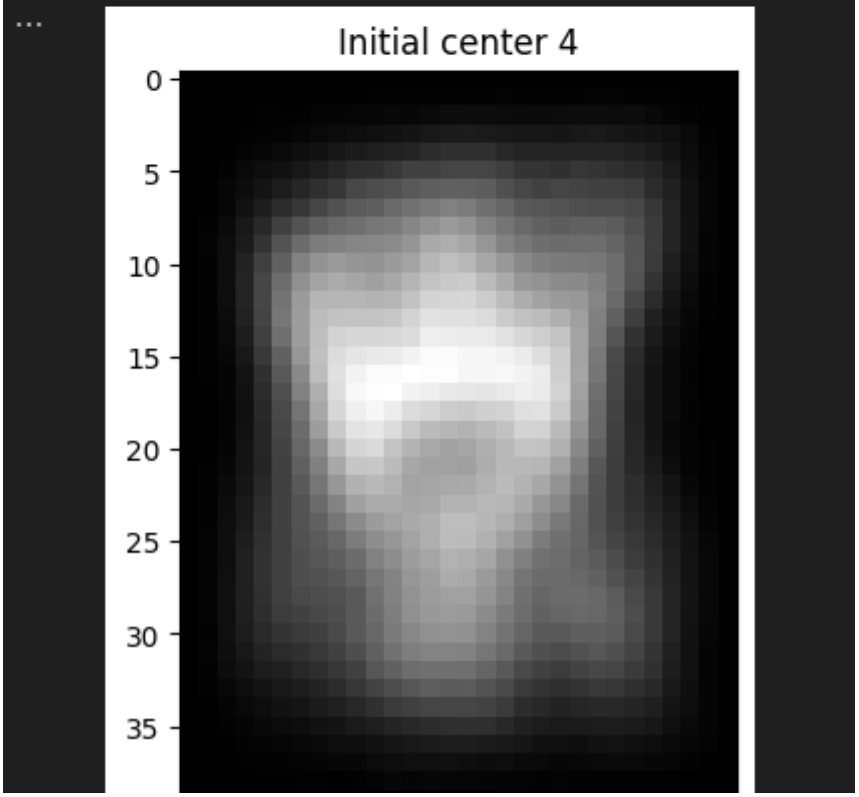


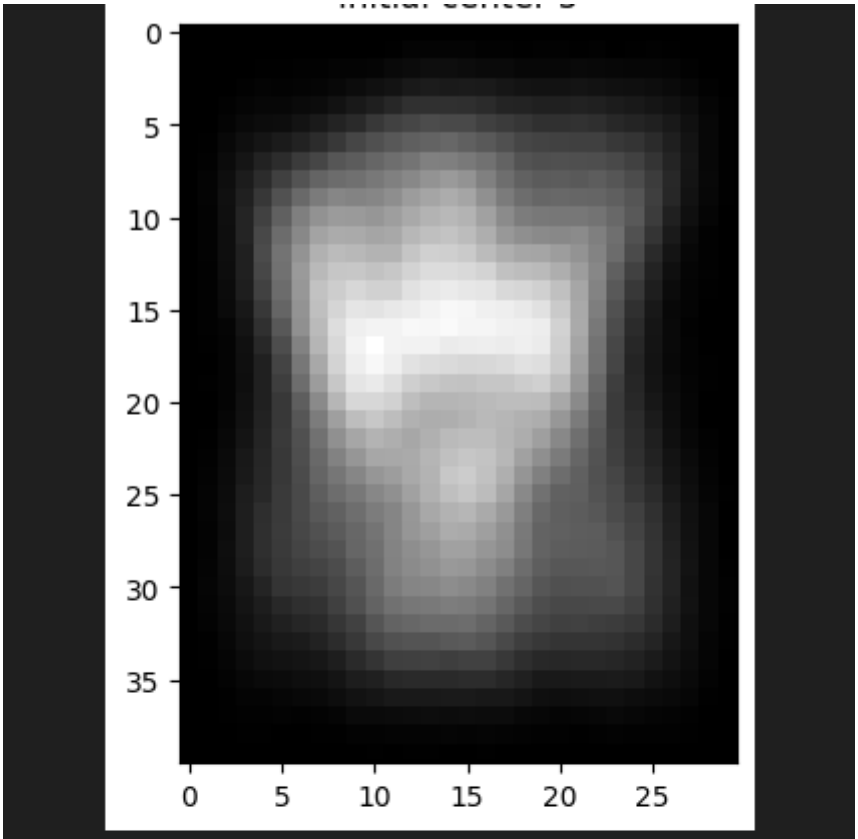
a.III:

شاخص های اولیه به صورت میانگین هر بخش

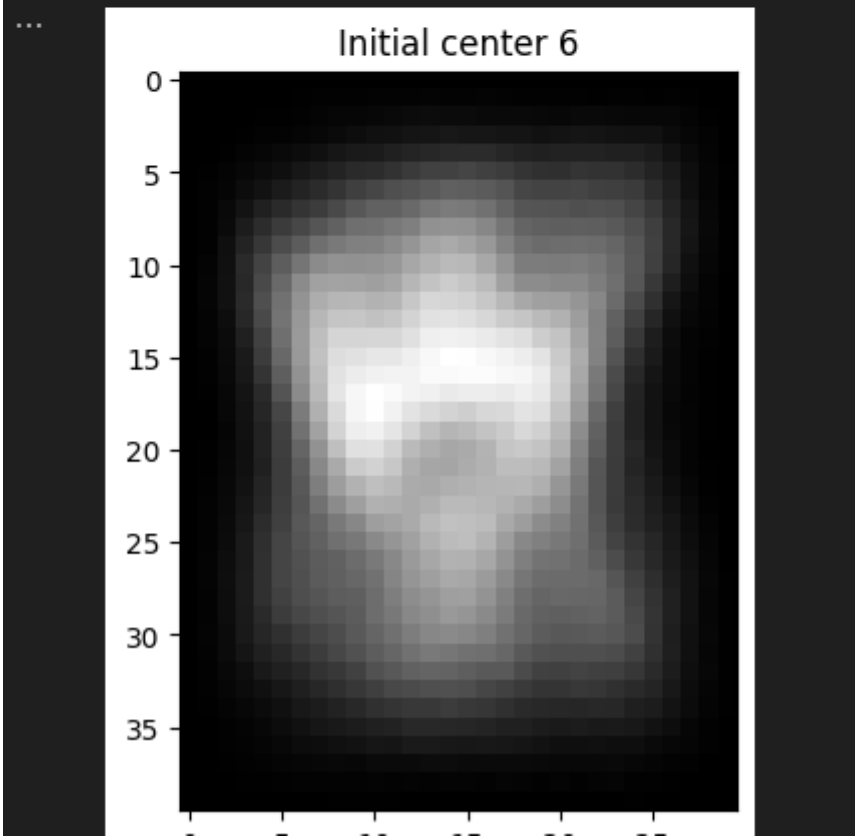


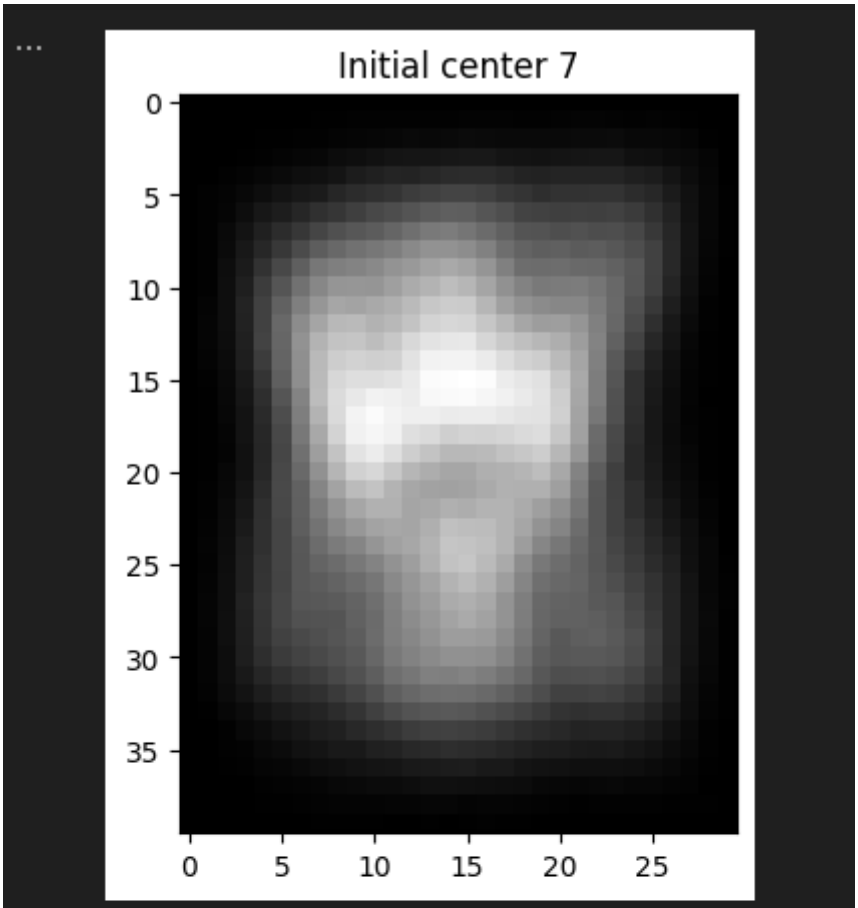

Initial center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25

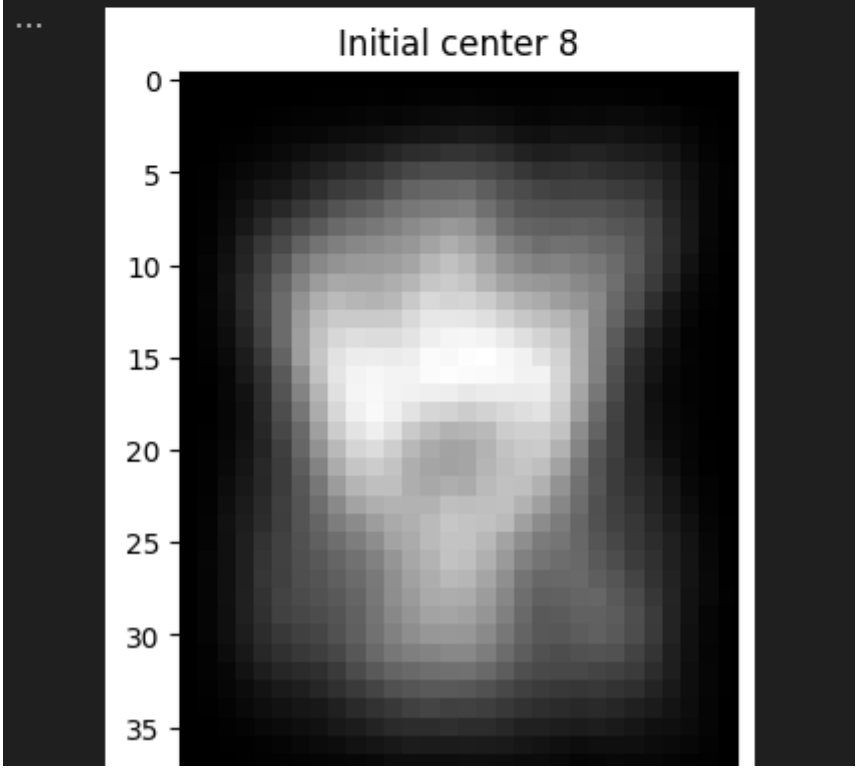

Initial center 4
0
5
10
15
20
25
30
35

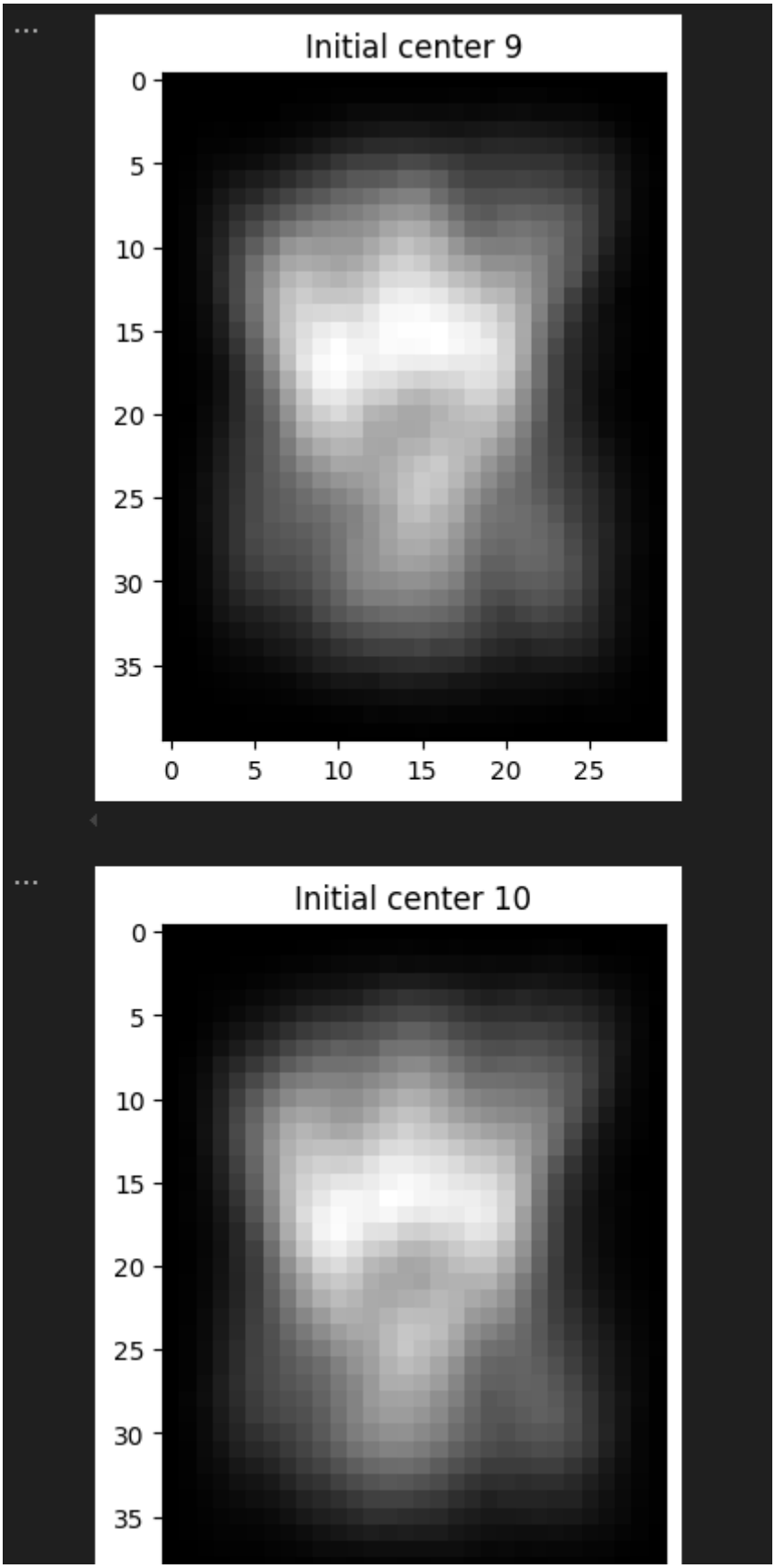

0
5
10
15
20
25
30
35
0
5
10
15
20
25

...

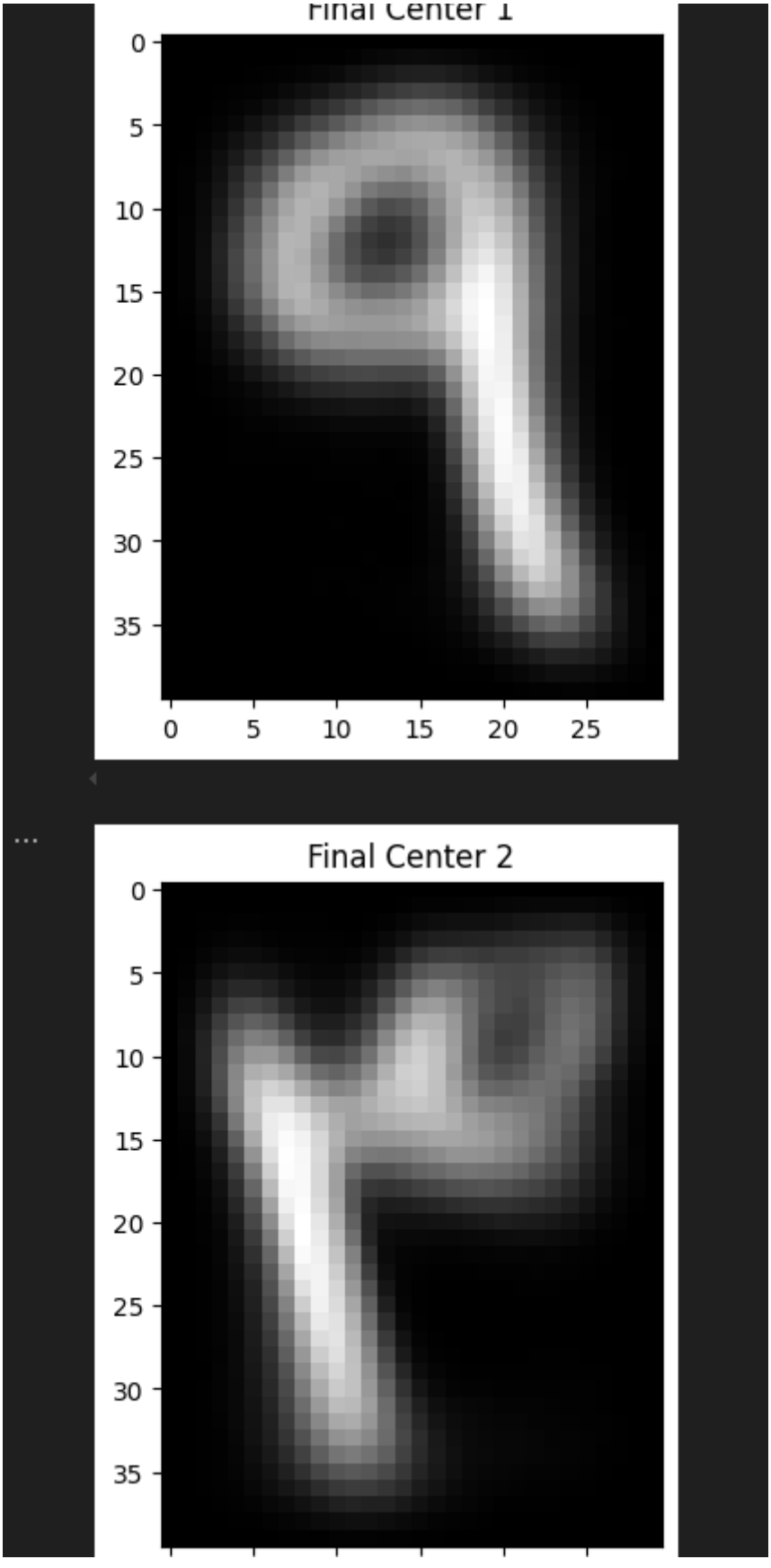

Initial center 6
0
5
10
15
20
25
30
35

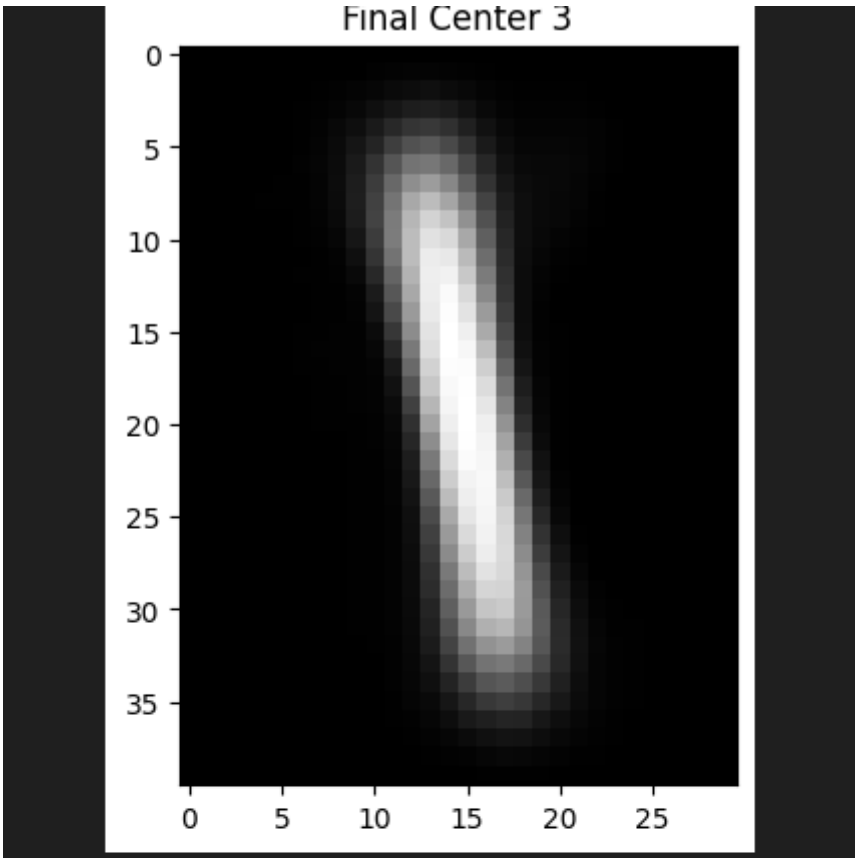

Initial center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Initial center 8
0
5
10
15
20
25
30
35

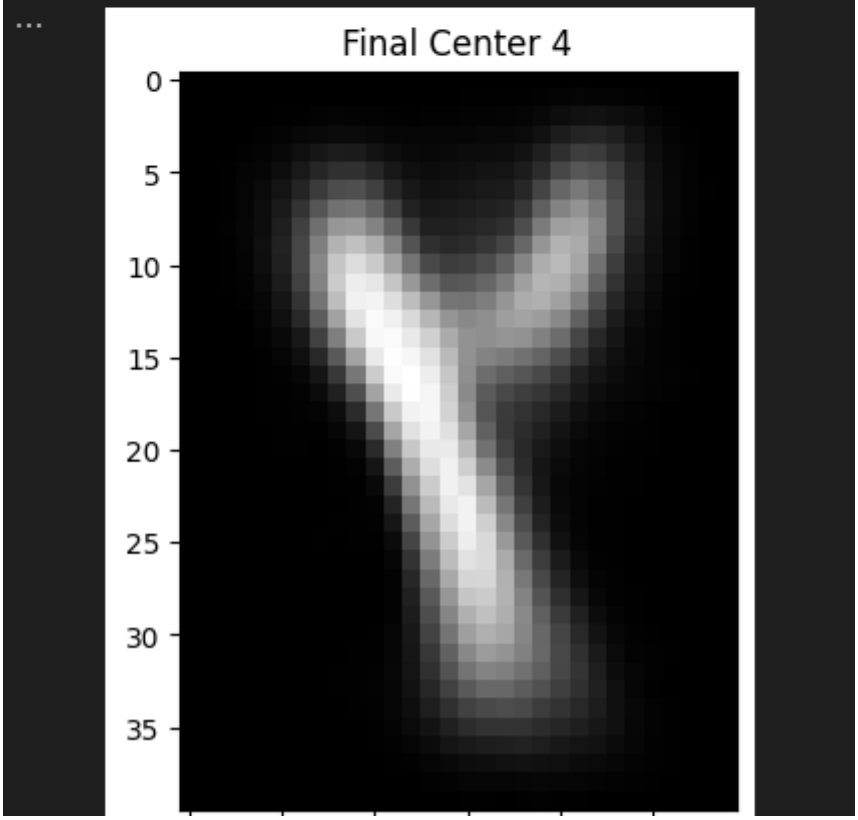

Initial center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial center 10
0
5
10
15
20
25
30
35

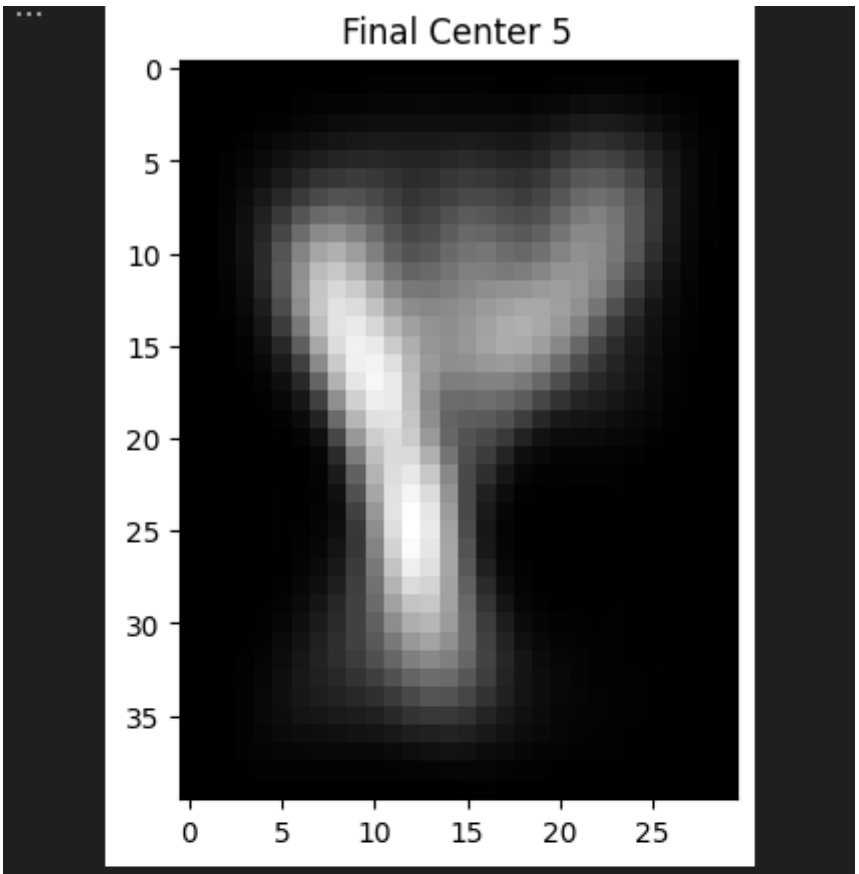
شاخص های نهایی

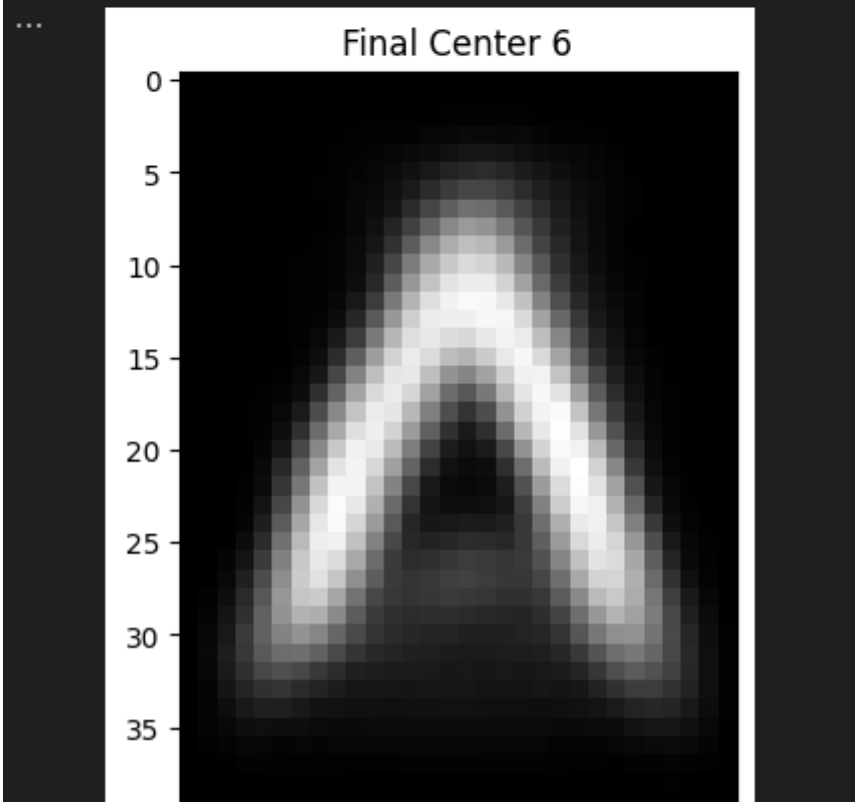


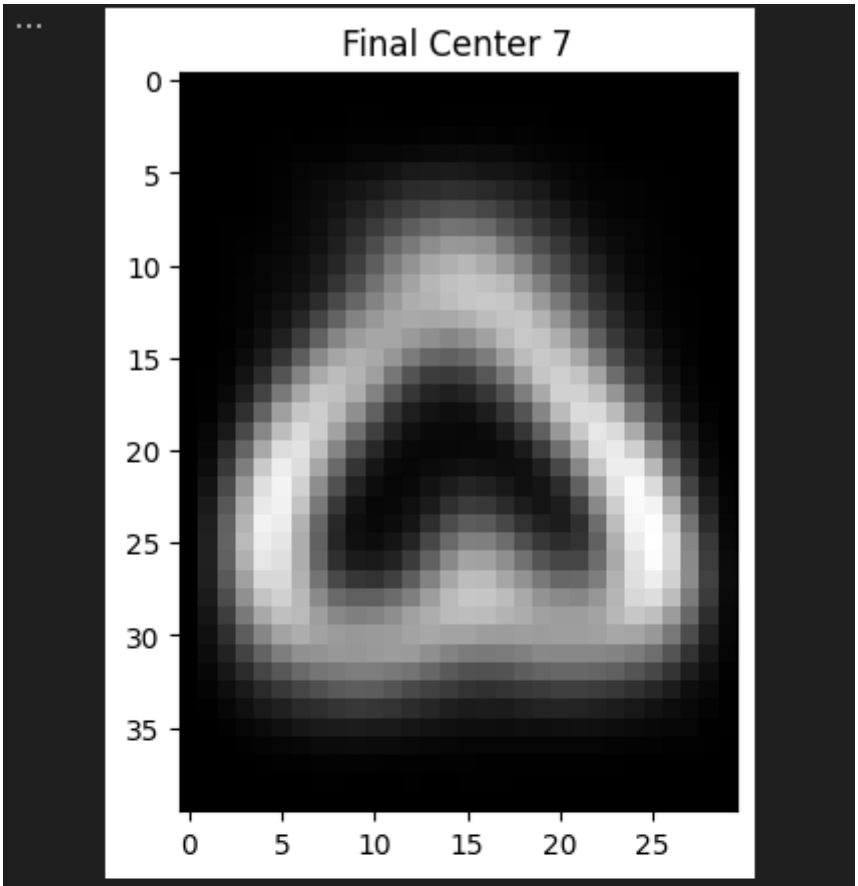

Final Center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25

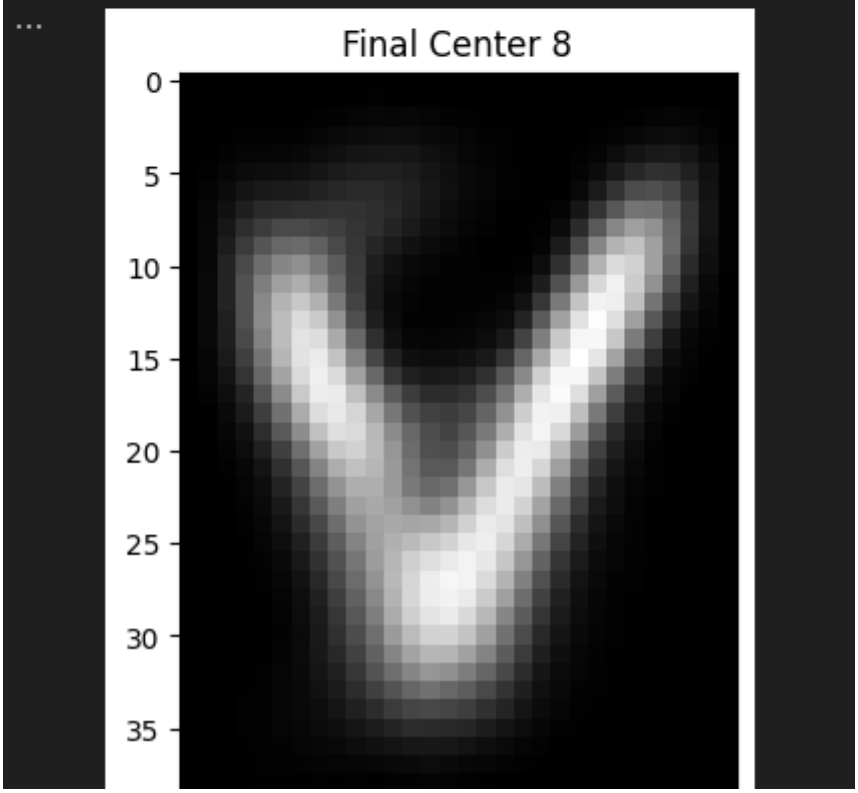
...

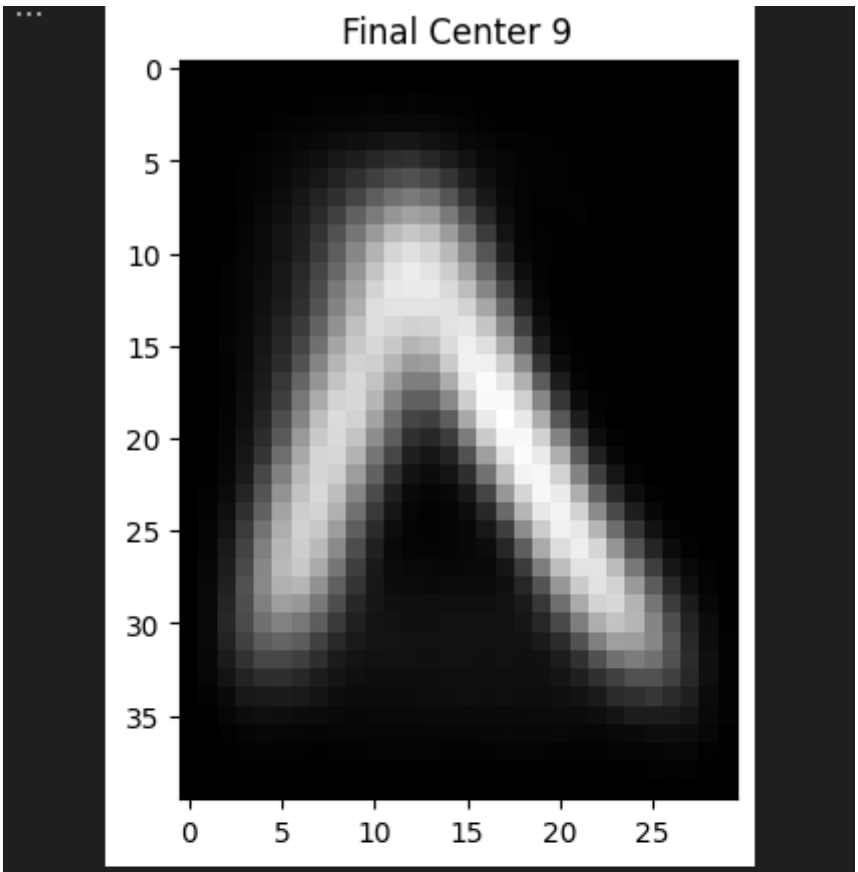

Final Center 4
0
5
10
15
20
25
30
35

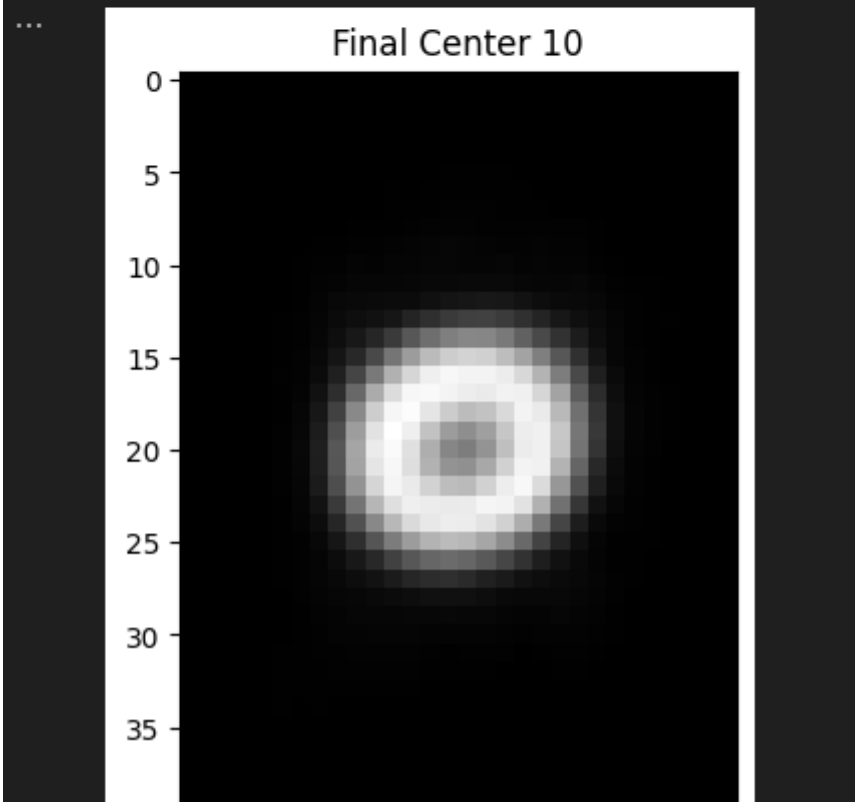

Final Center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 6
0
5
10
15
20
25
30
35

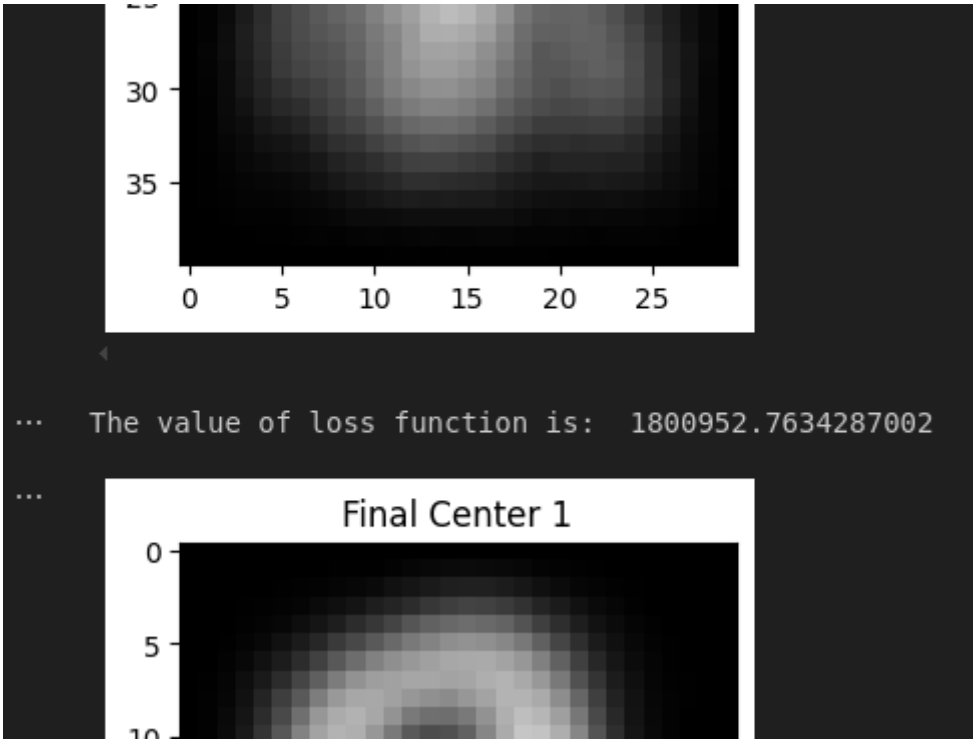

Final Center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 8
0
5
10
15
20
25
30
35


Final Center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25

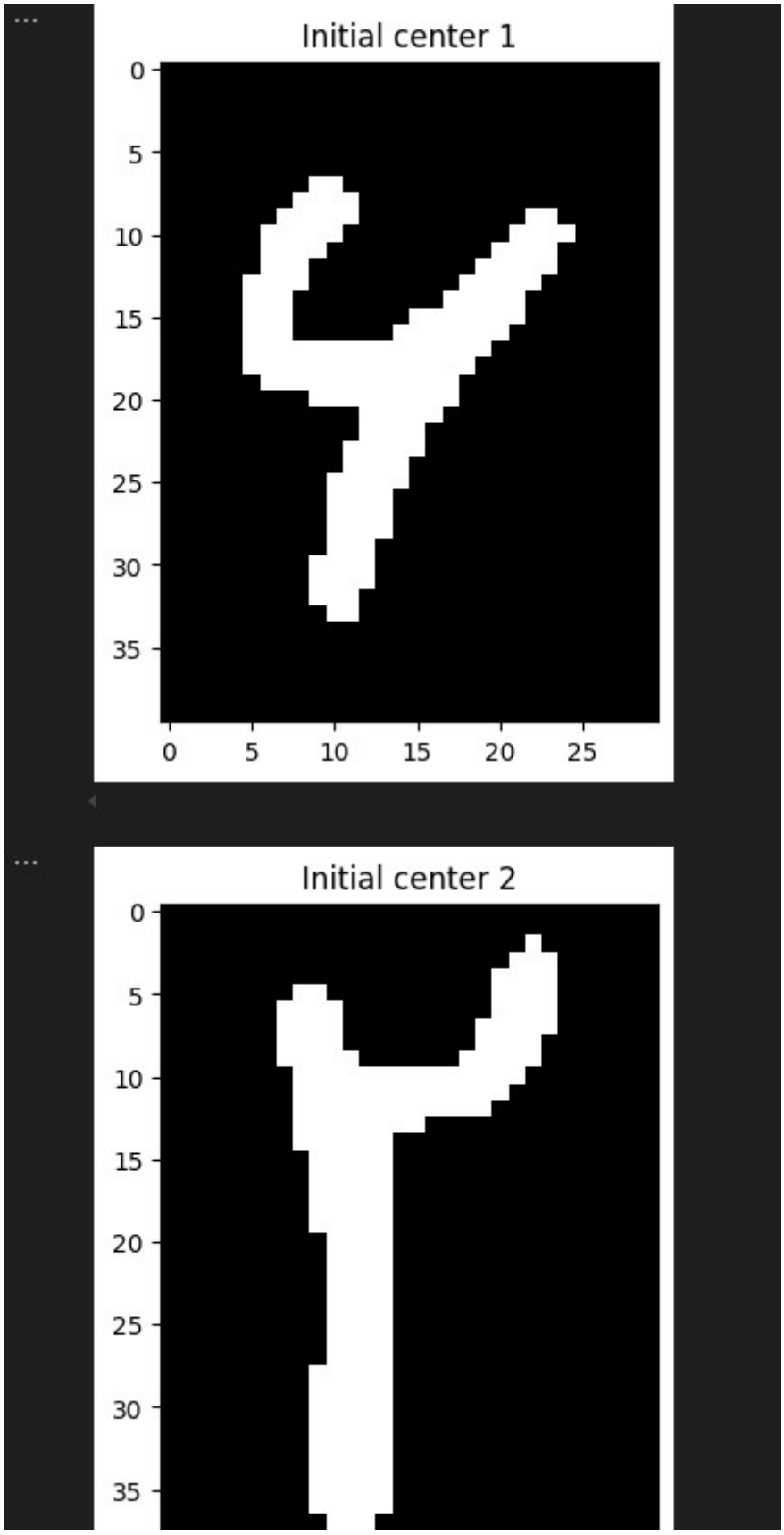

Final Center 10
0
5
10
15
20
25
30
35

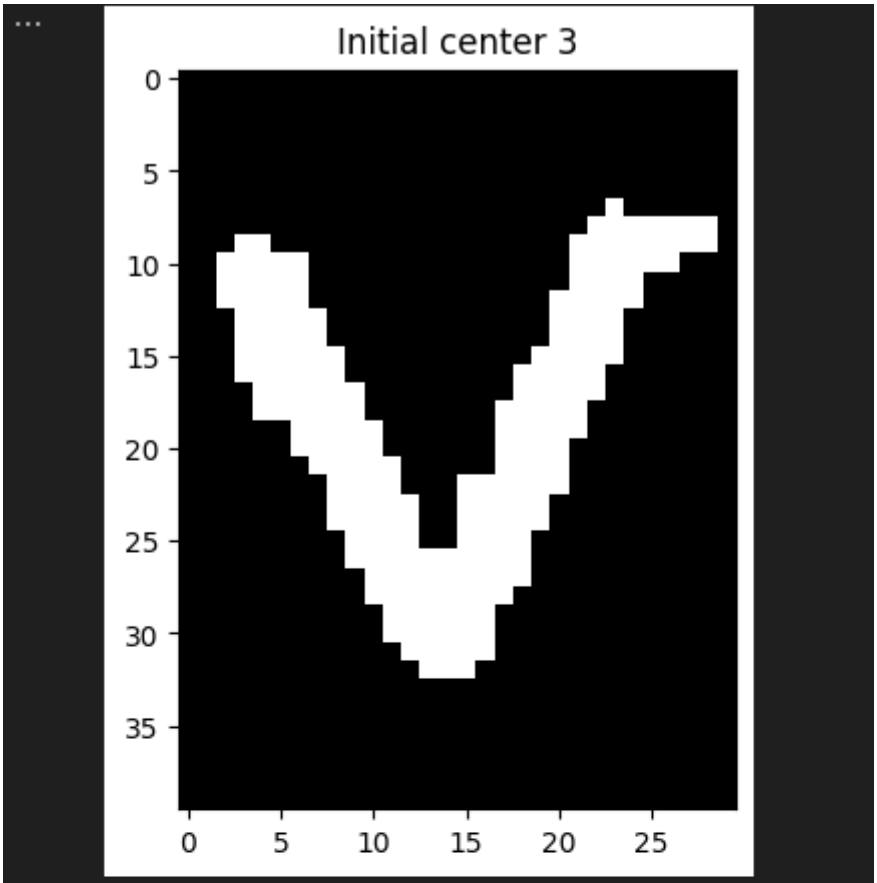
مقدار تابع هزینه

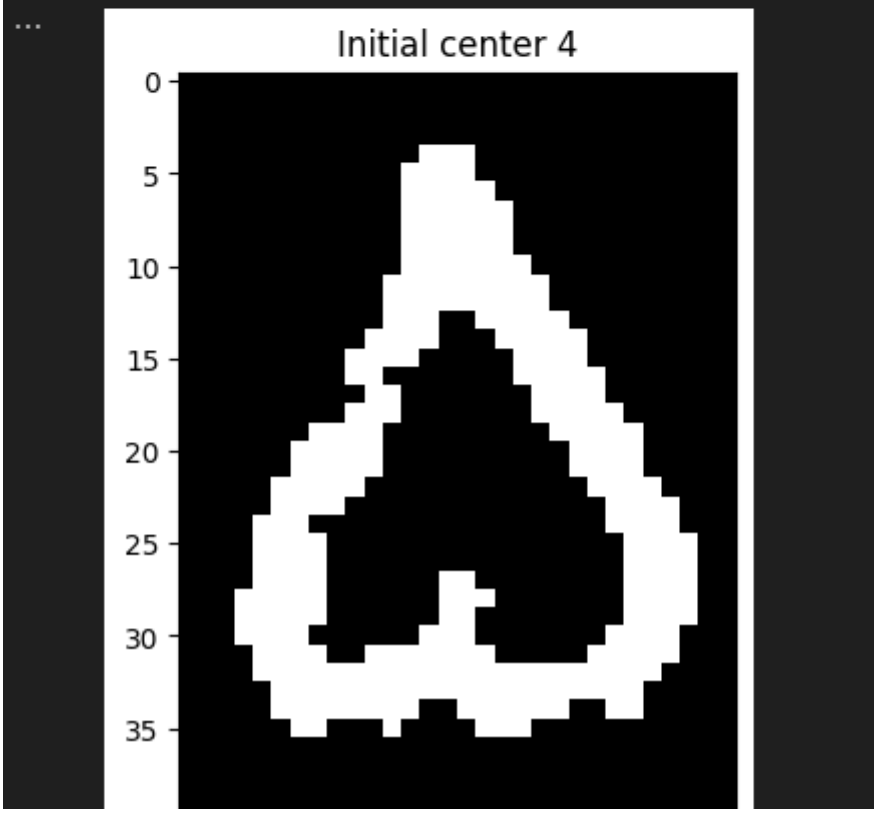


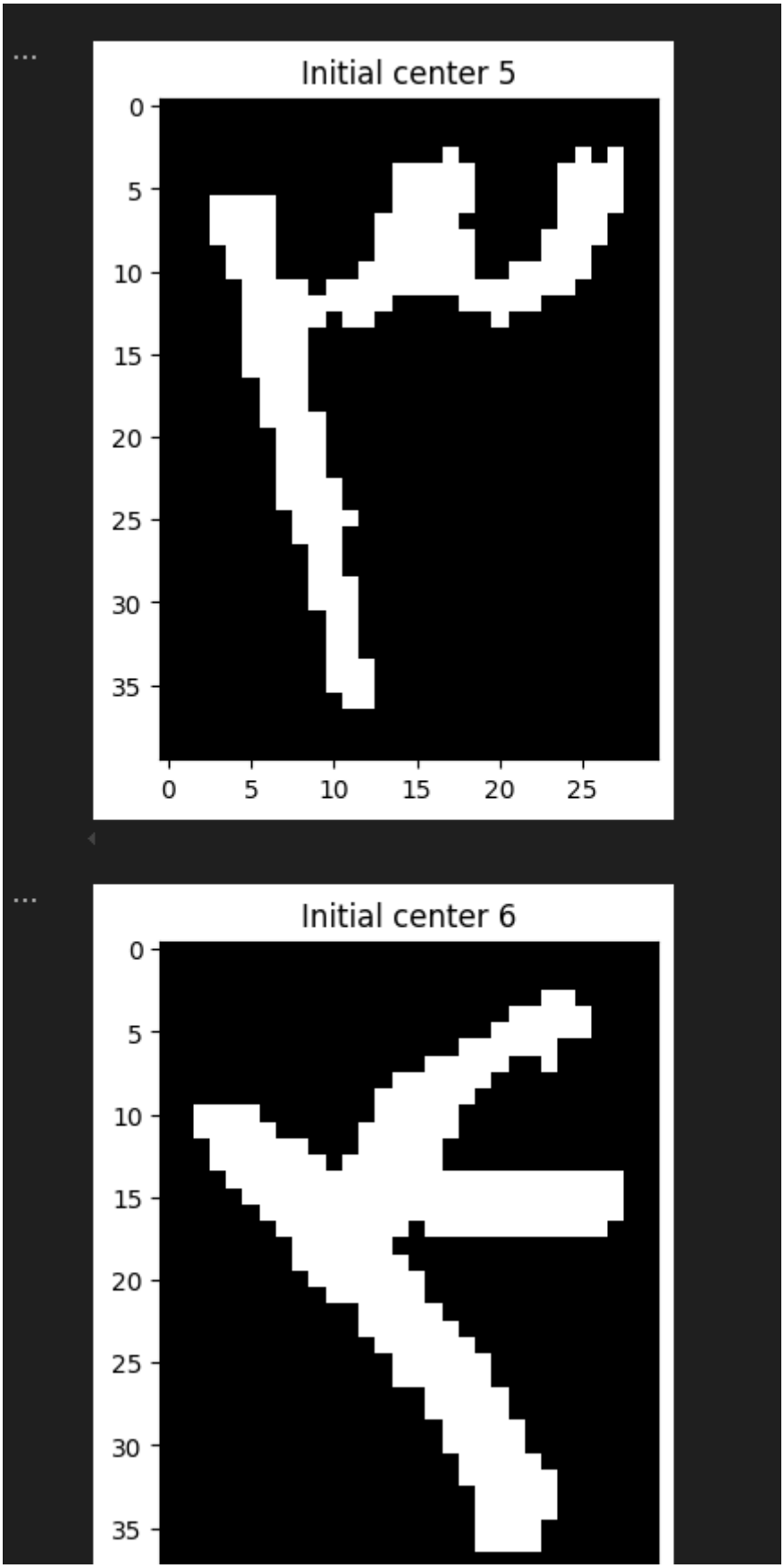
a.IV:

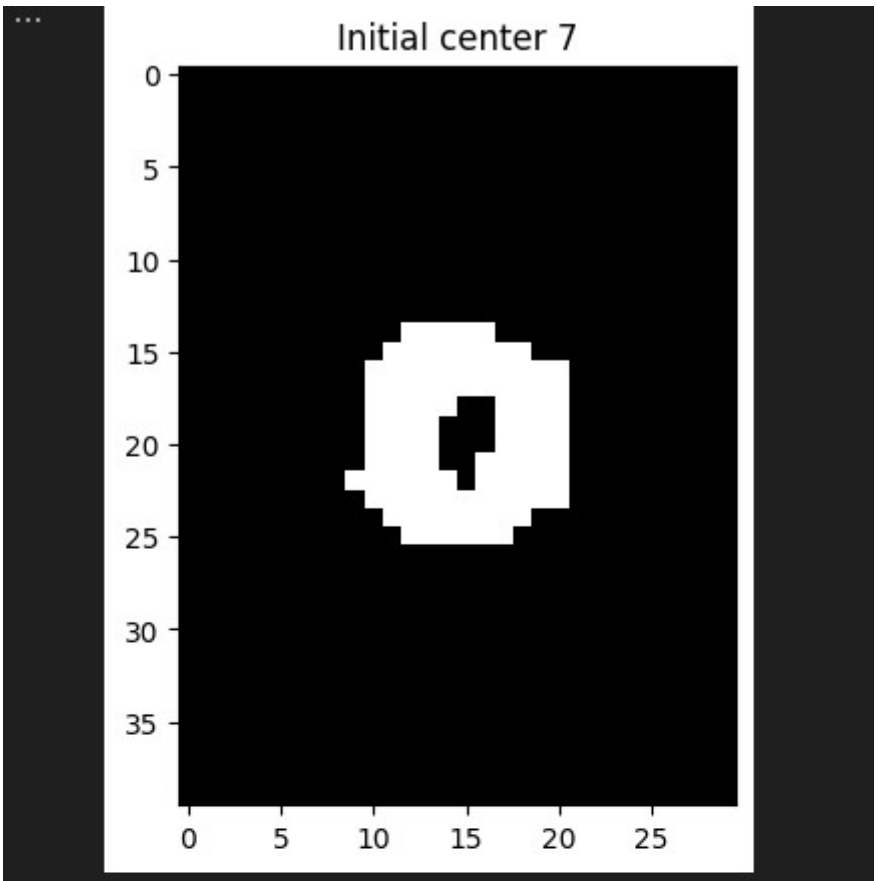
شاخص های اولیه مناسب انتخاب شده از بین ۱۰۰ داده اول

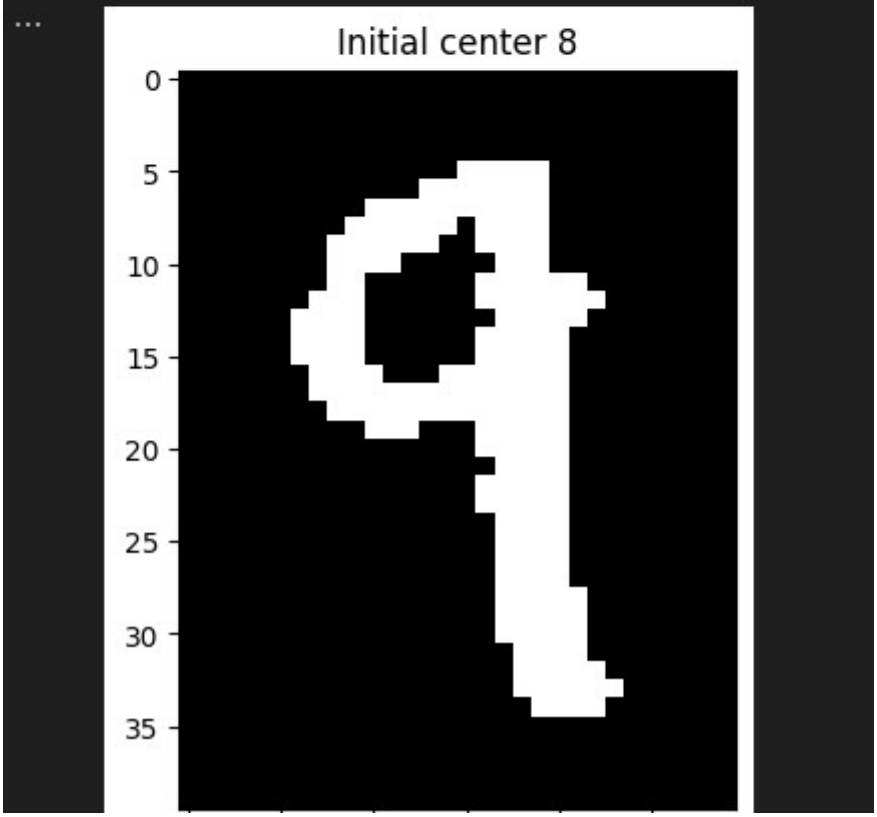


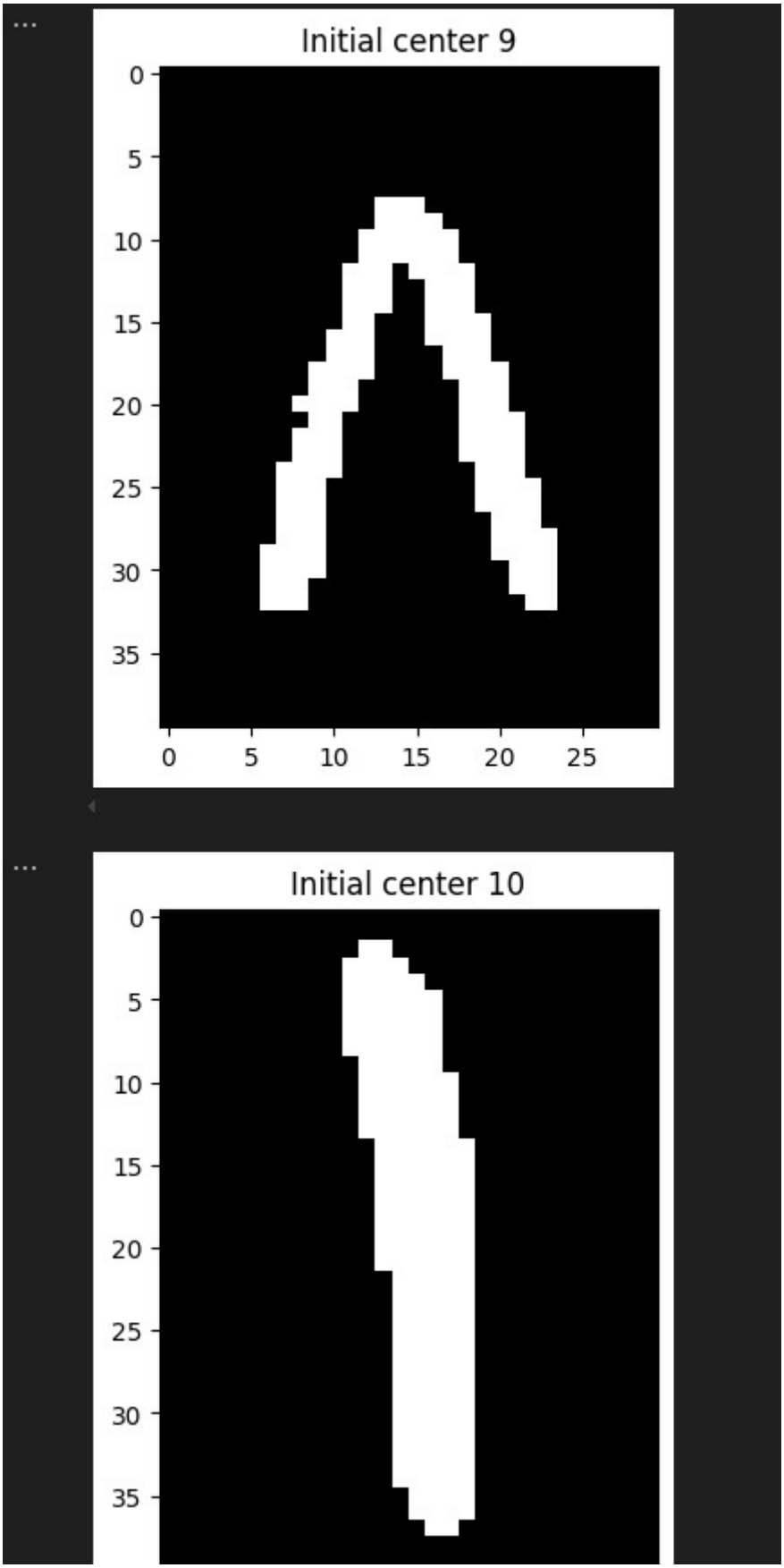

Initial center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Initial center 4
0
5
10
15
20
25
30
35

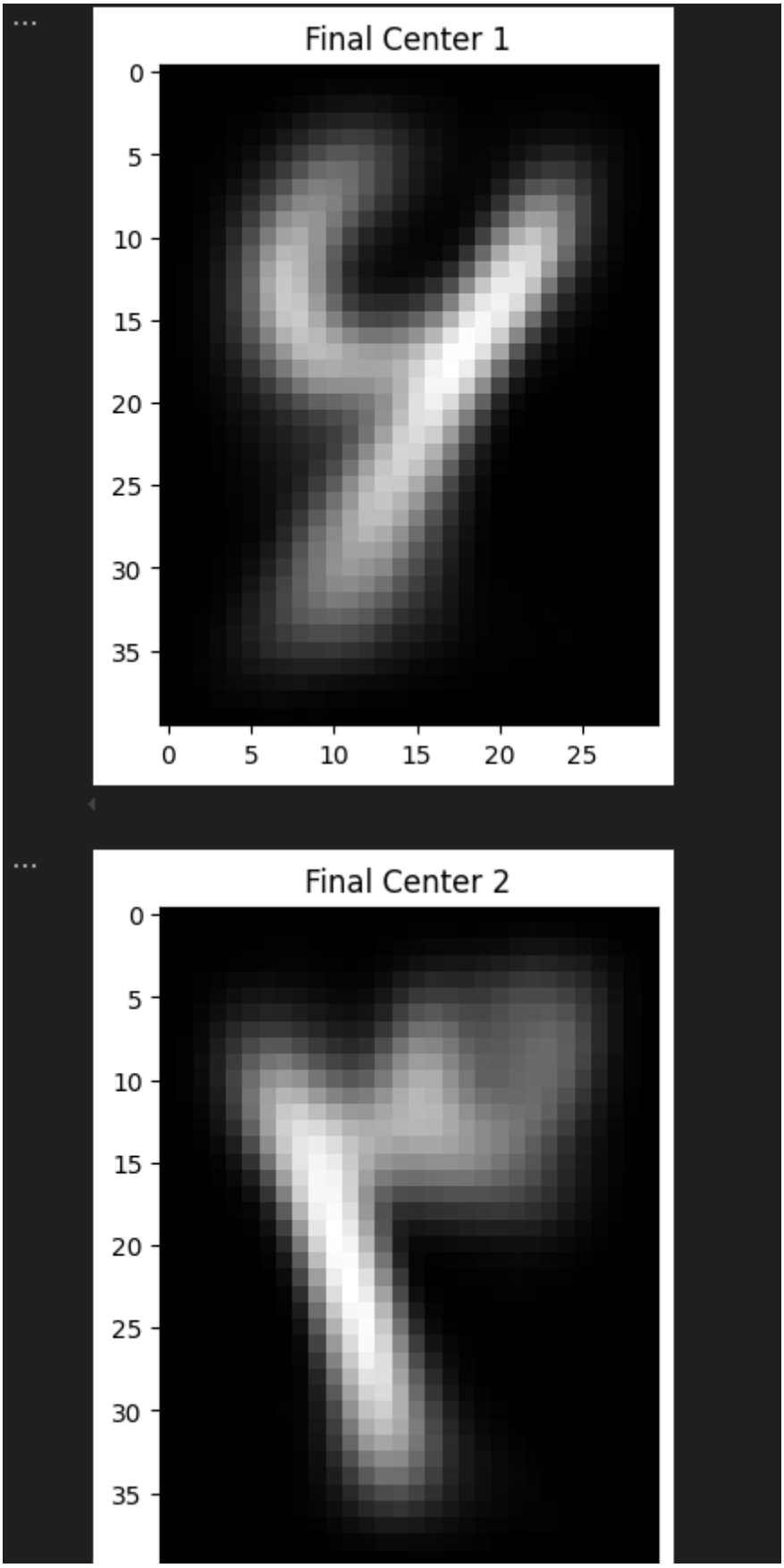

Initial center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial center 6
0
5
10
15
20
25
30
35

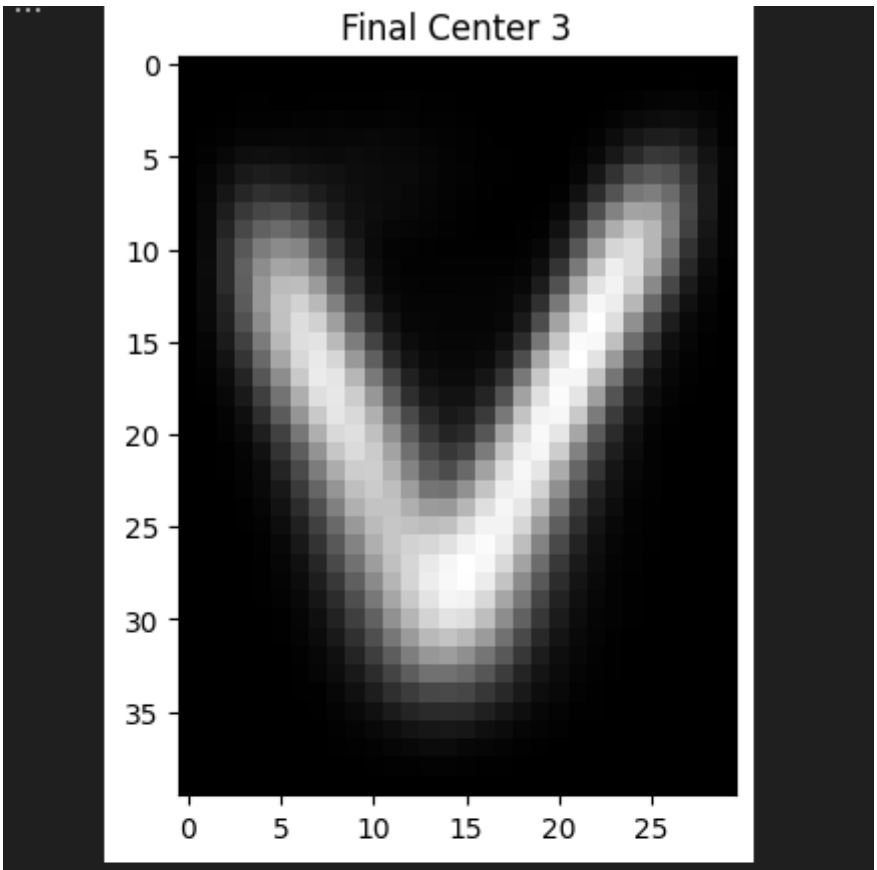

Initial center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25

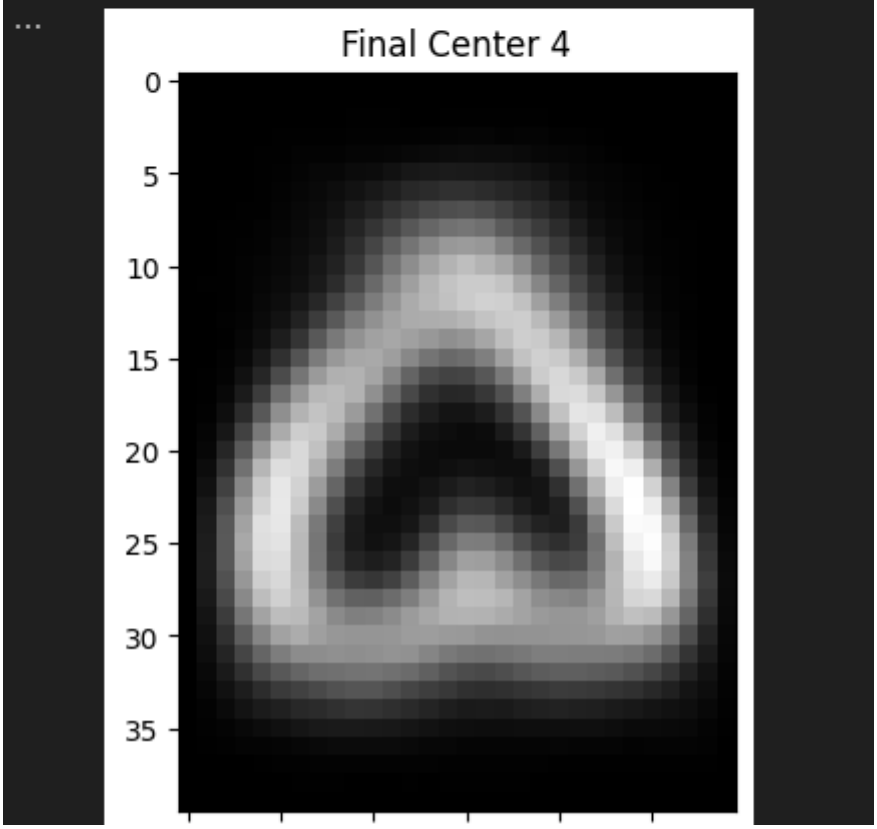

Initial center 8
0
5
10
15
20
25
30
35

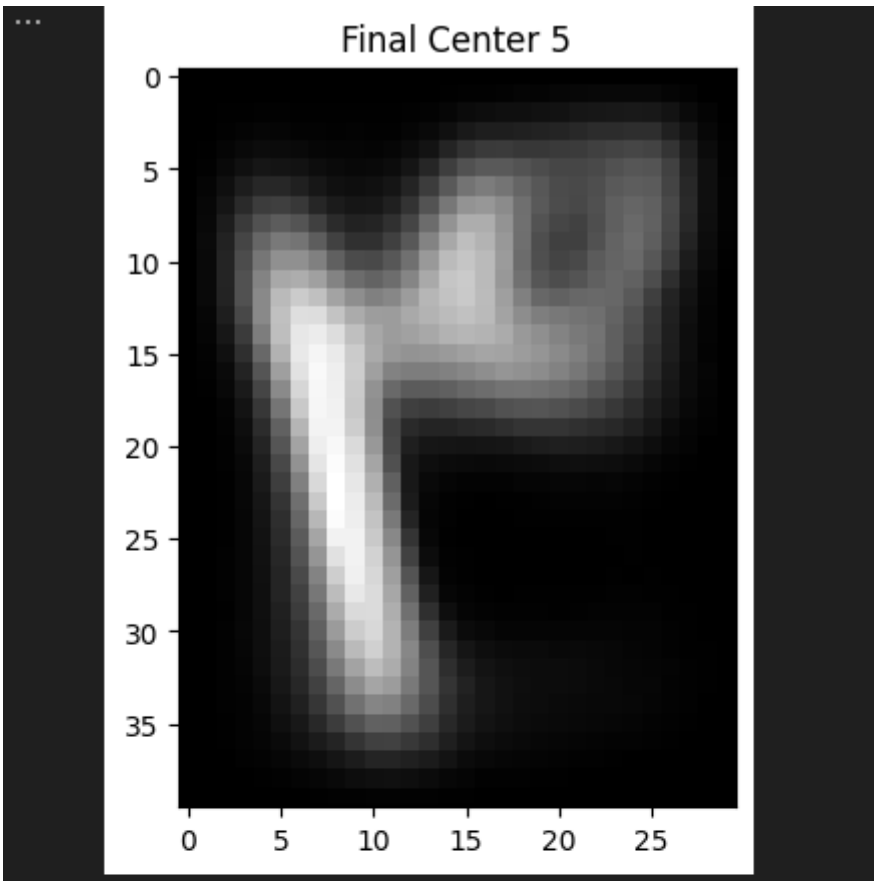

Initial center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial center 10
0
5
10
15
20
25
30
35

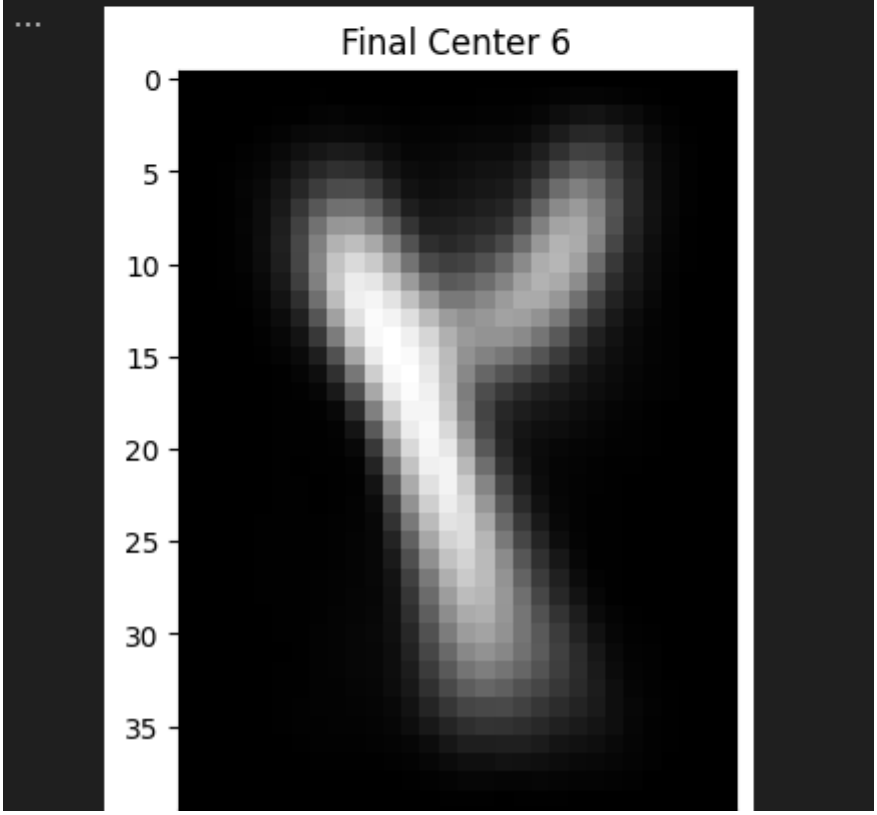
شاخص های نهایی

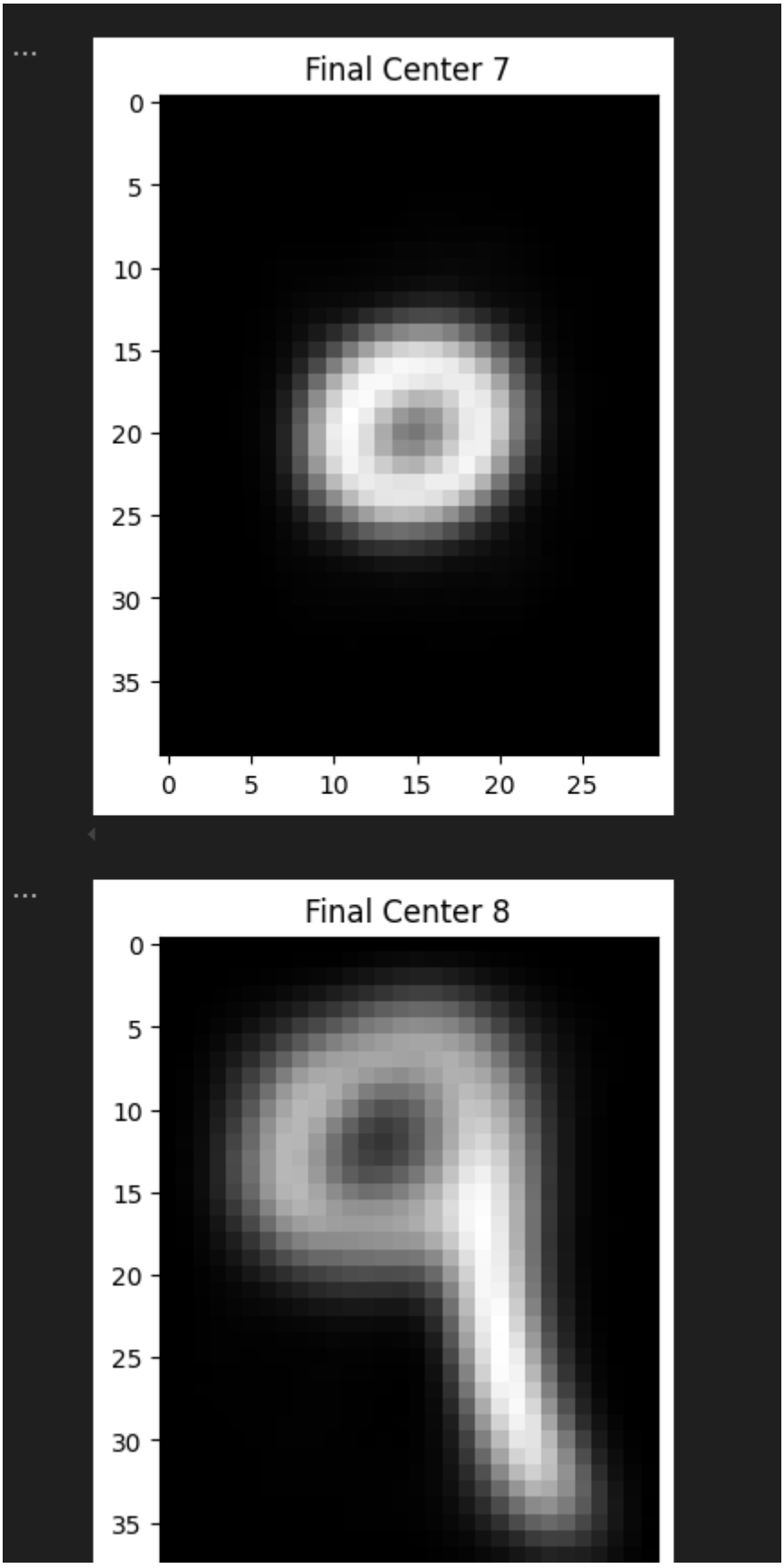


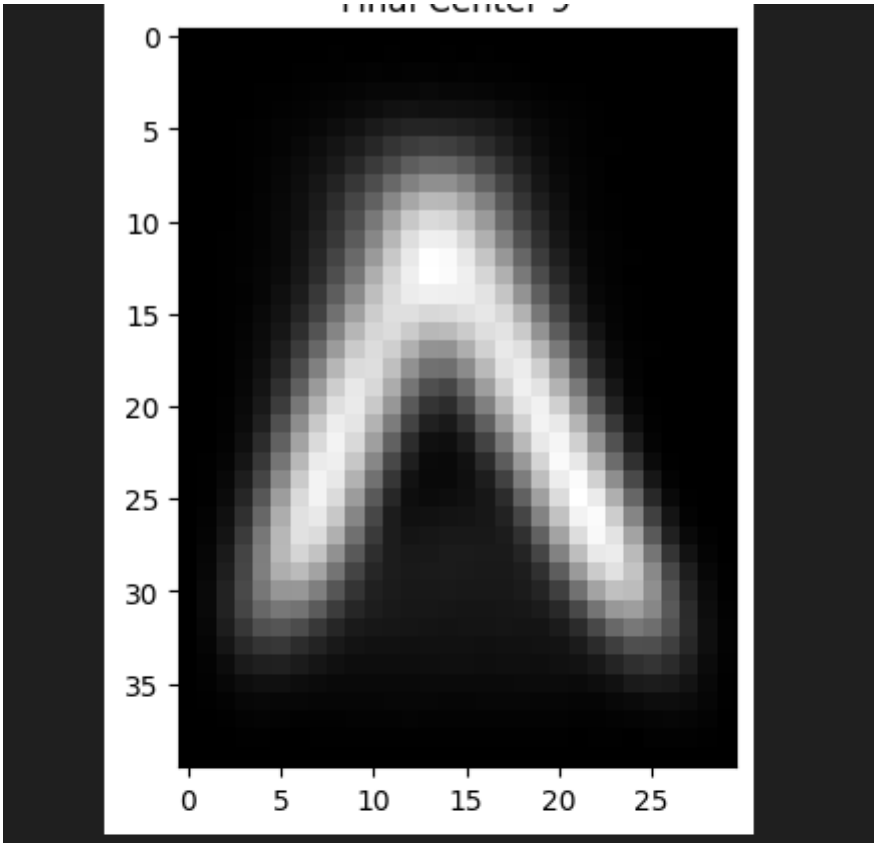

Final Center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 4
0
5
10
15
20
25
30
35

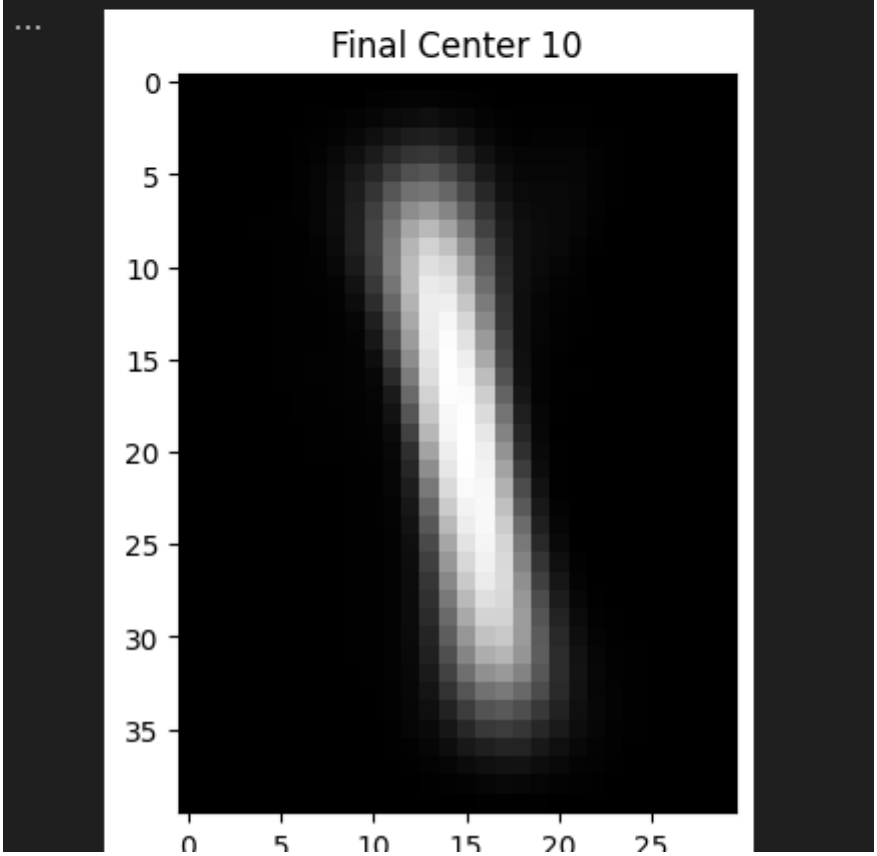

Final Center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 6
0
5
10
15
20
25
30
35

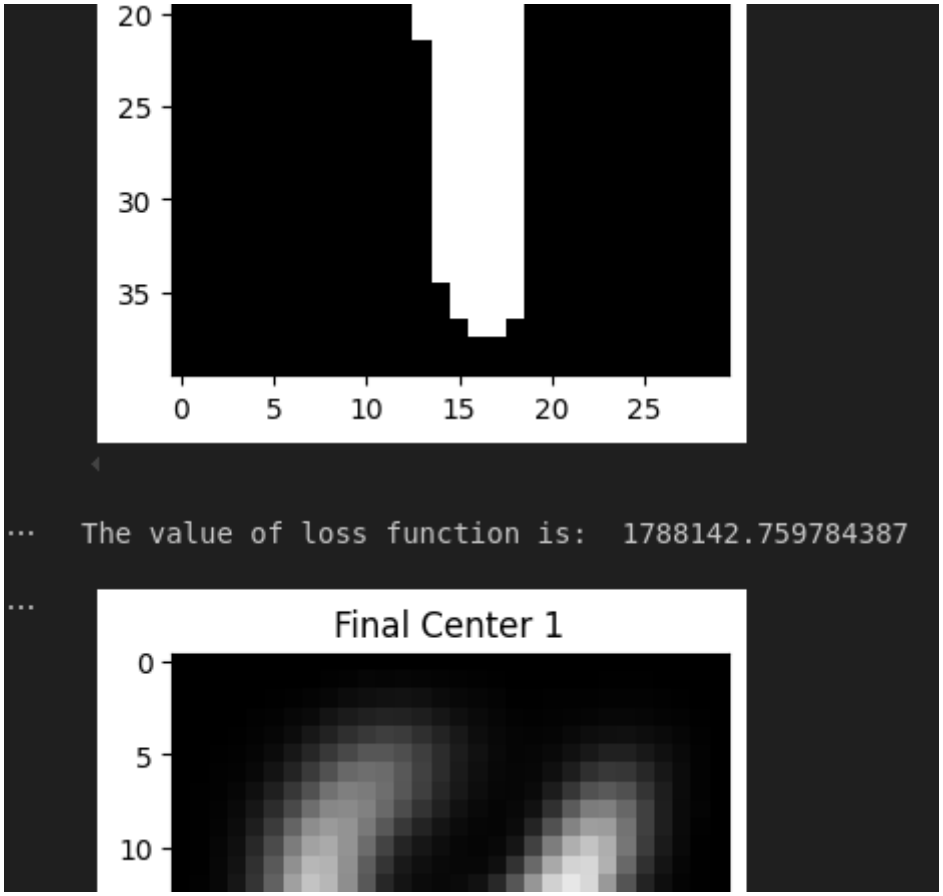

Final Center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 8
0
5
10
15
20
25
30
35


0
5
10
15
20
25
30
35
0
5
10
15
20
25

...

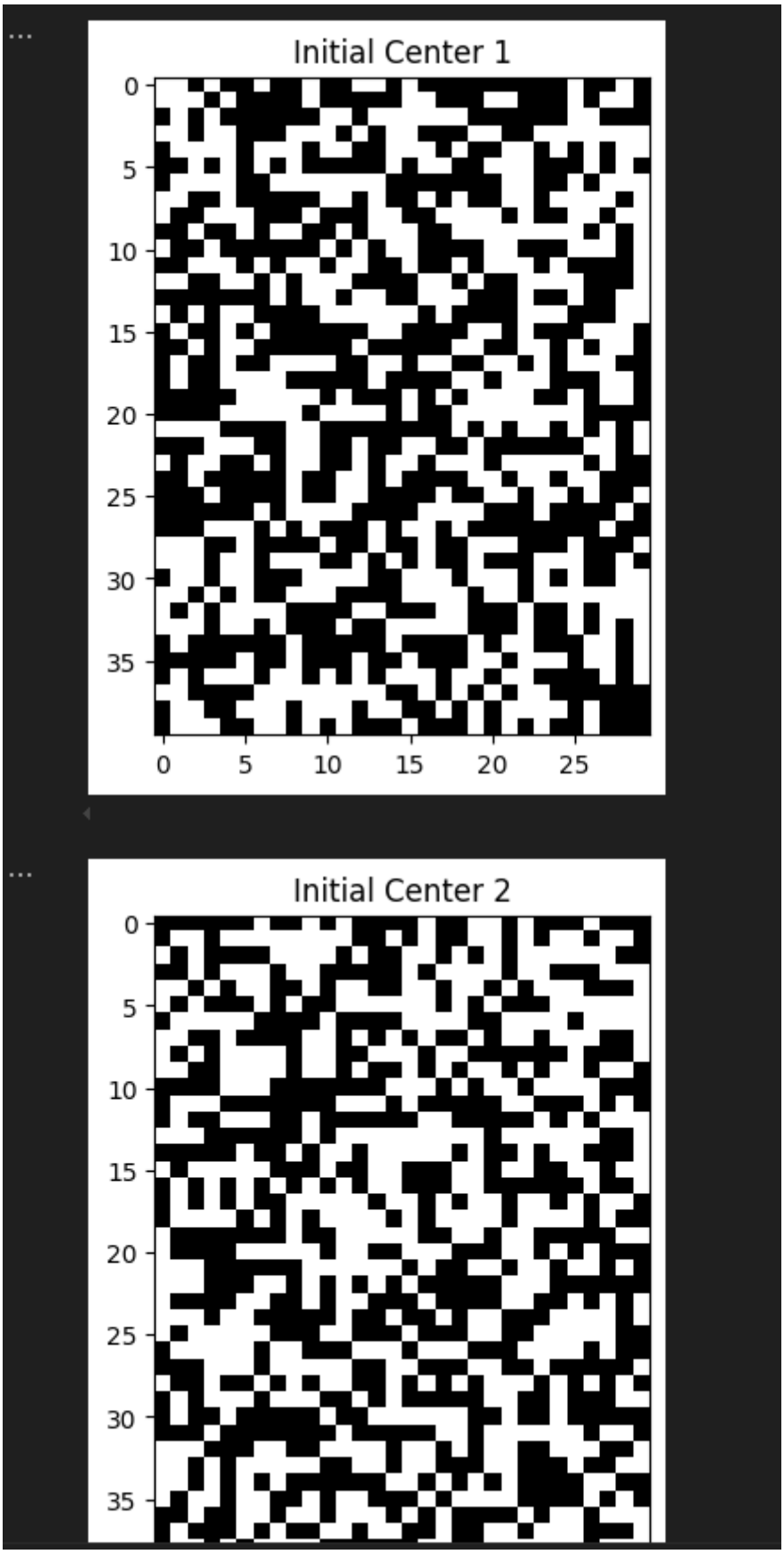

Final Center 10
0
5
10
15
20
25
30
35
0
5
10
15
20
25

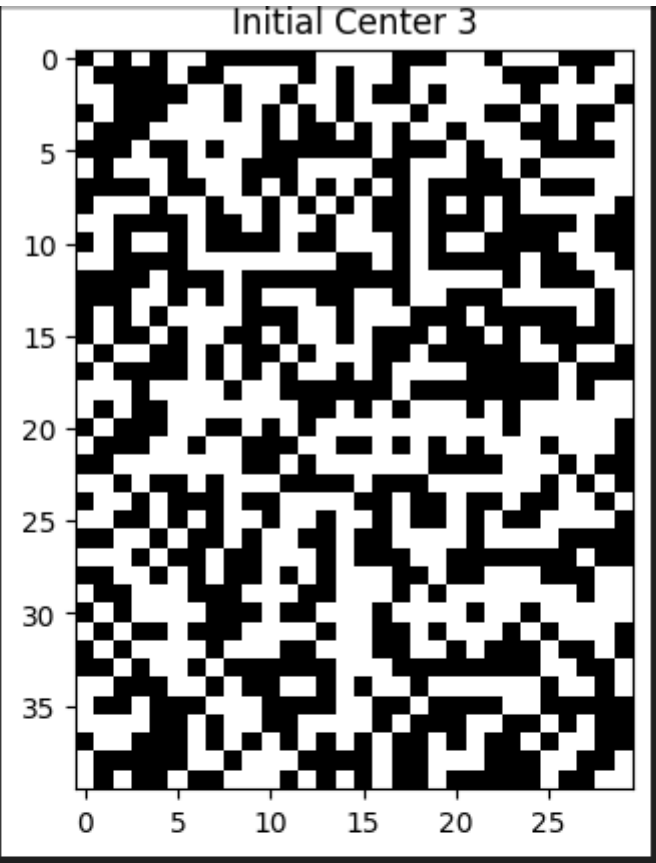
مقدار تابع هزینه



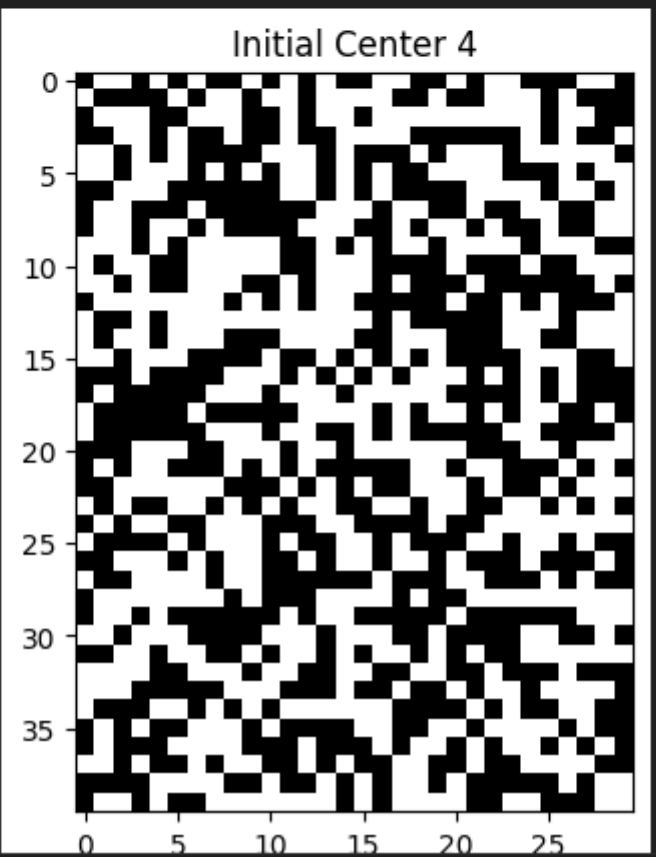
b.I:

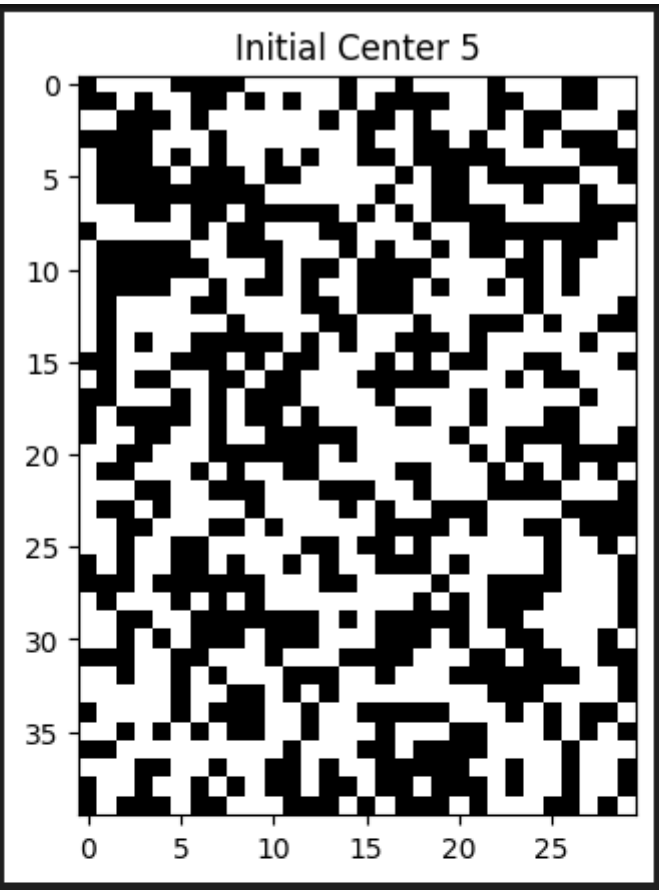
شاخص های اولیه تصادفی ساخته شده

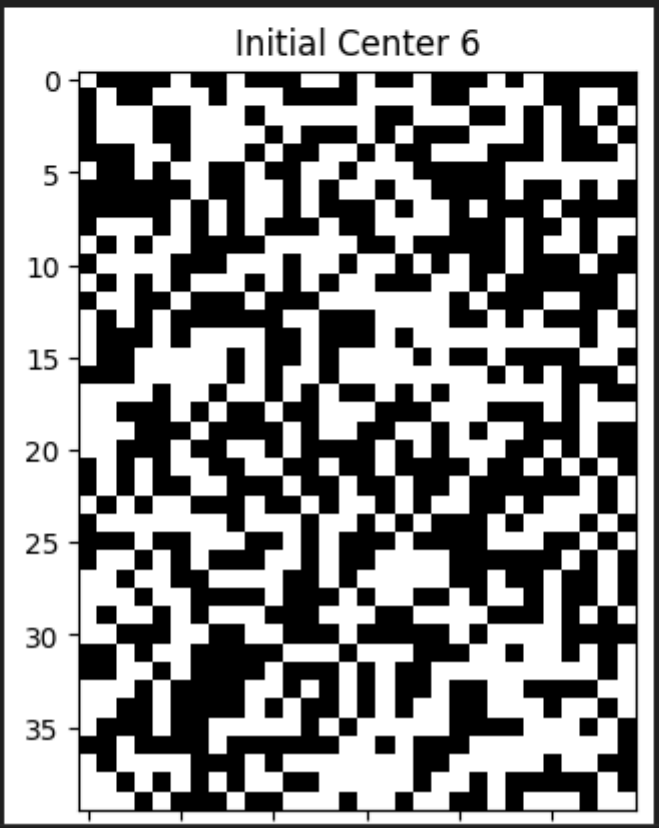


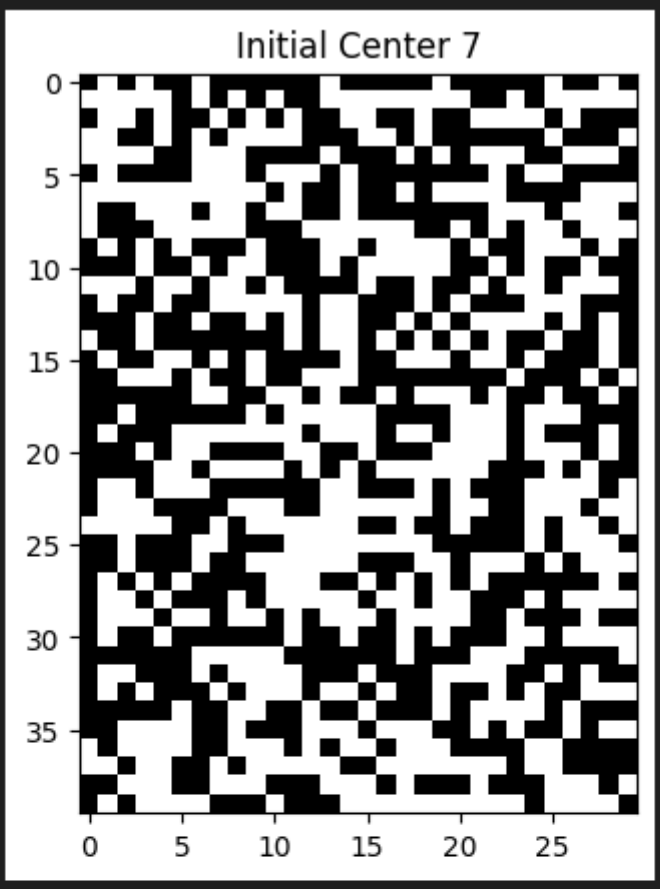

Initial Center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25

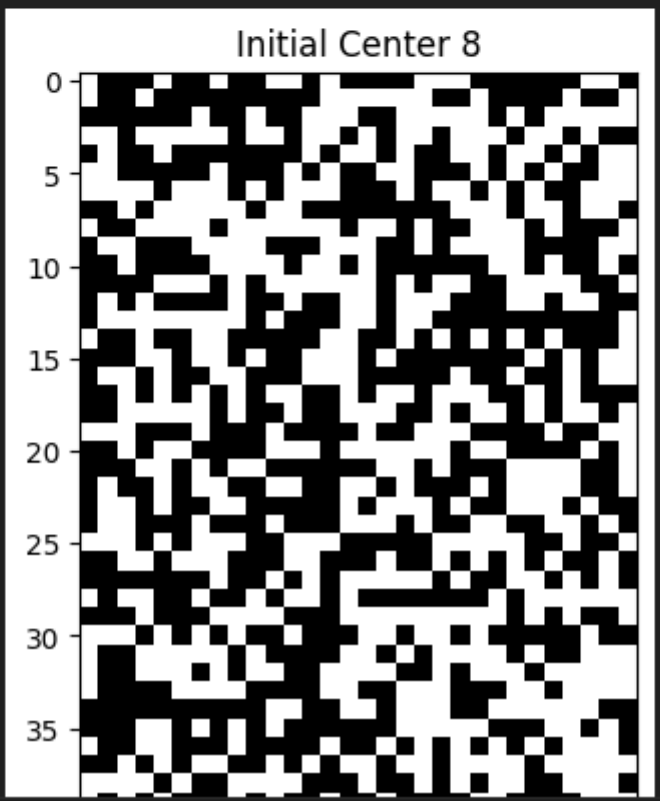
...

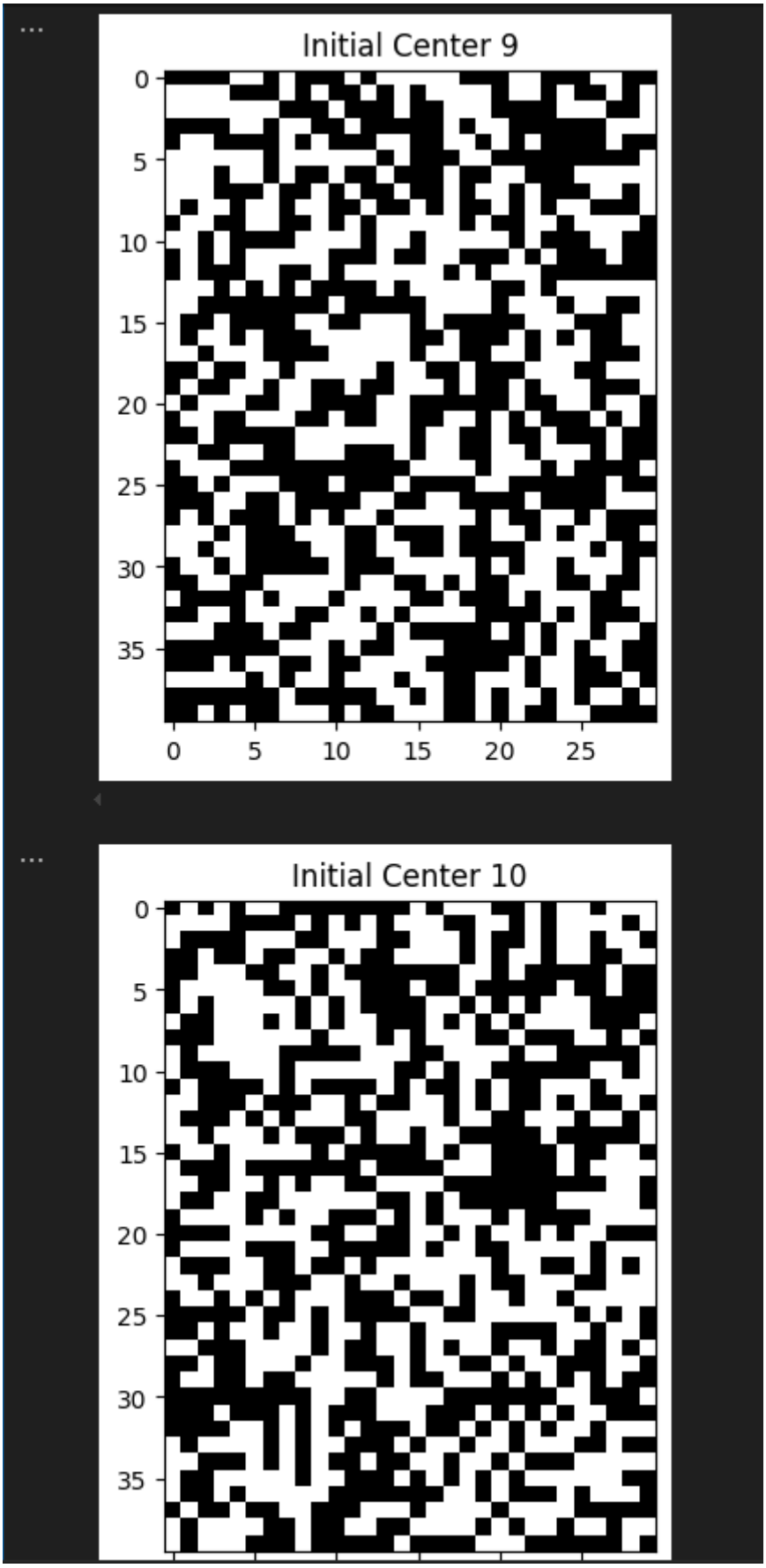

Initial Center 4
0
5
10
15
20
25
30
35
0
5
10
15
20
25

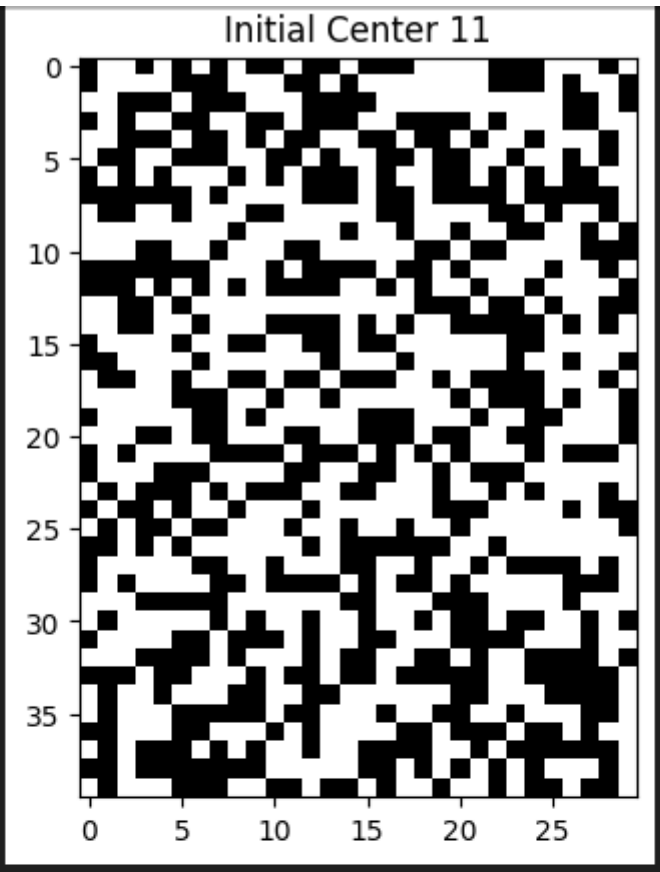

Initial Center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25

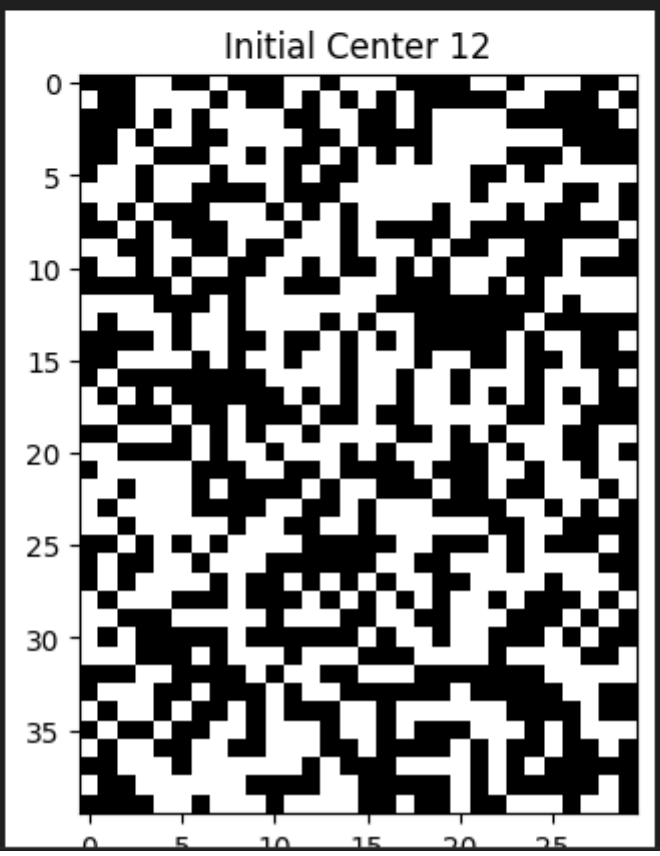

Initial Center 6
0
5
10
15
20
25
30
35

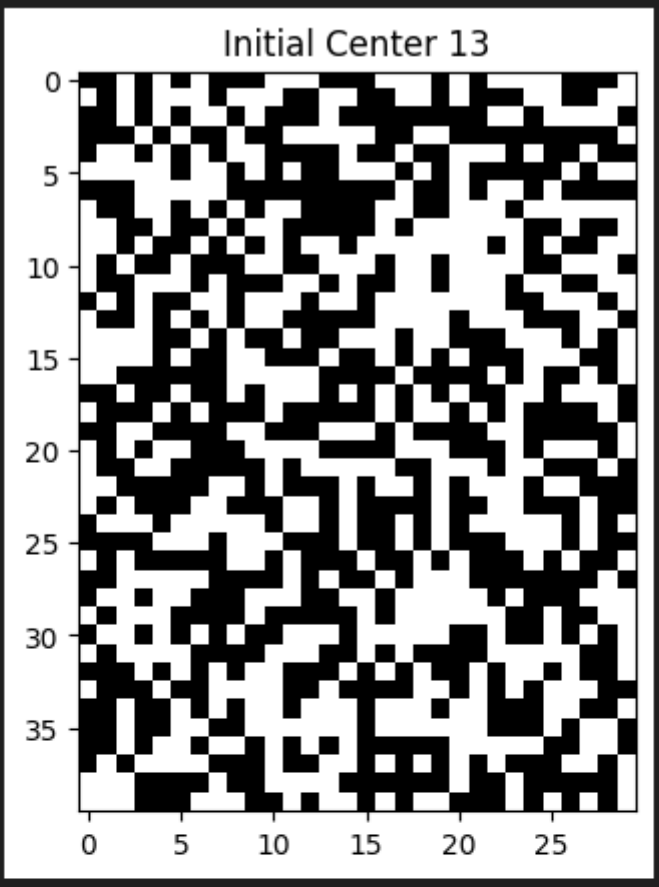

Initial Center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Initial Center 8
0
5
10
15
20
25
30
35


Initial Center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial Center 10
0
5
10
15
20
25
30
35

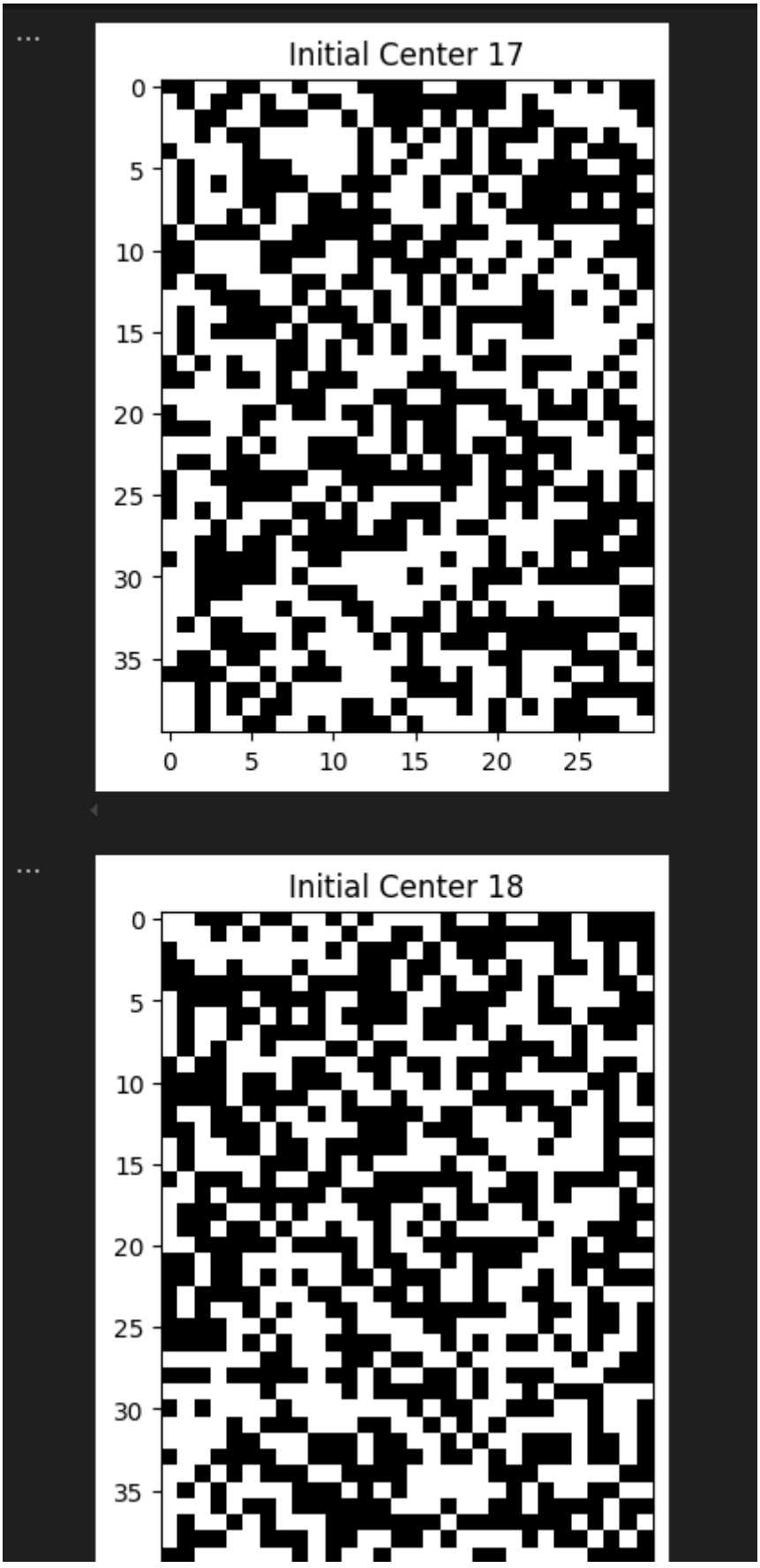

Initial Center 11
0
5
10
15
20
25
30
35
0
5
10
15
20
25

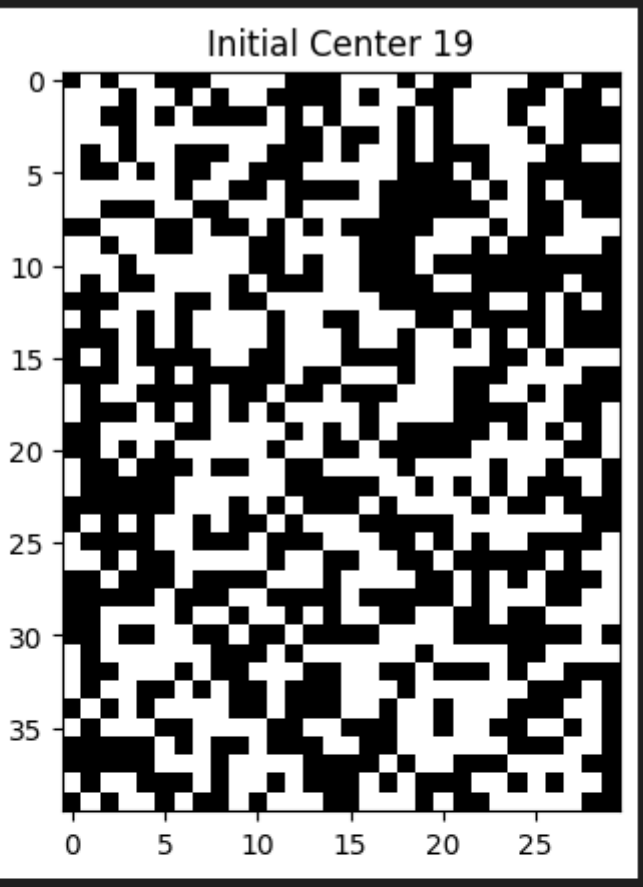

Initial Center 12
0
5
10
15
20
25
30
35

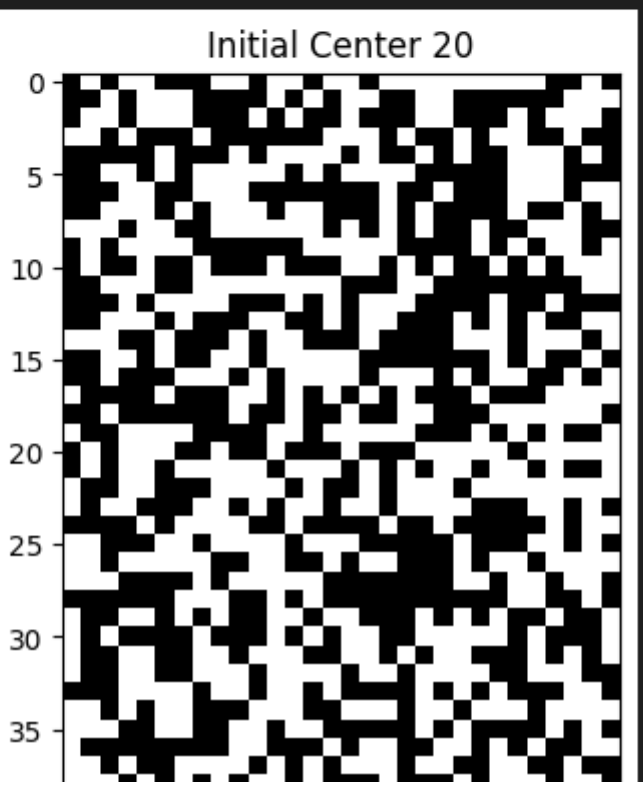

Initial Center 13
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Initial Center 14
0
5
10
15
20
25
30
35

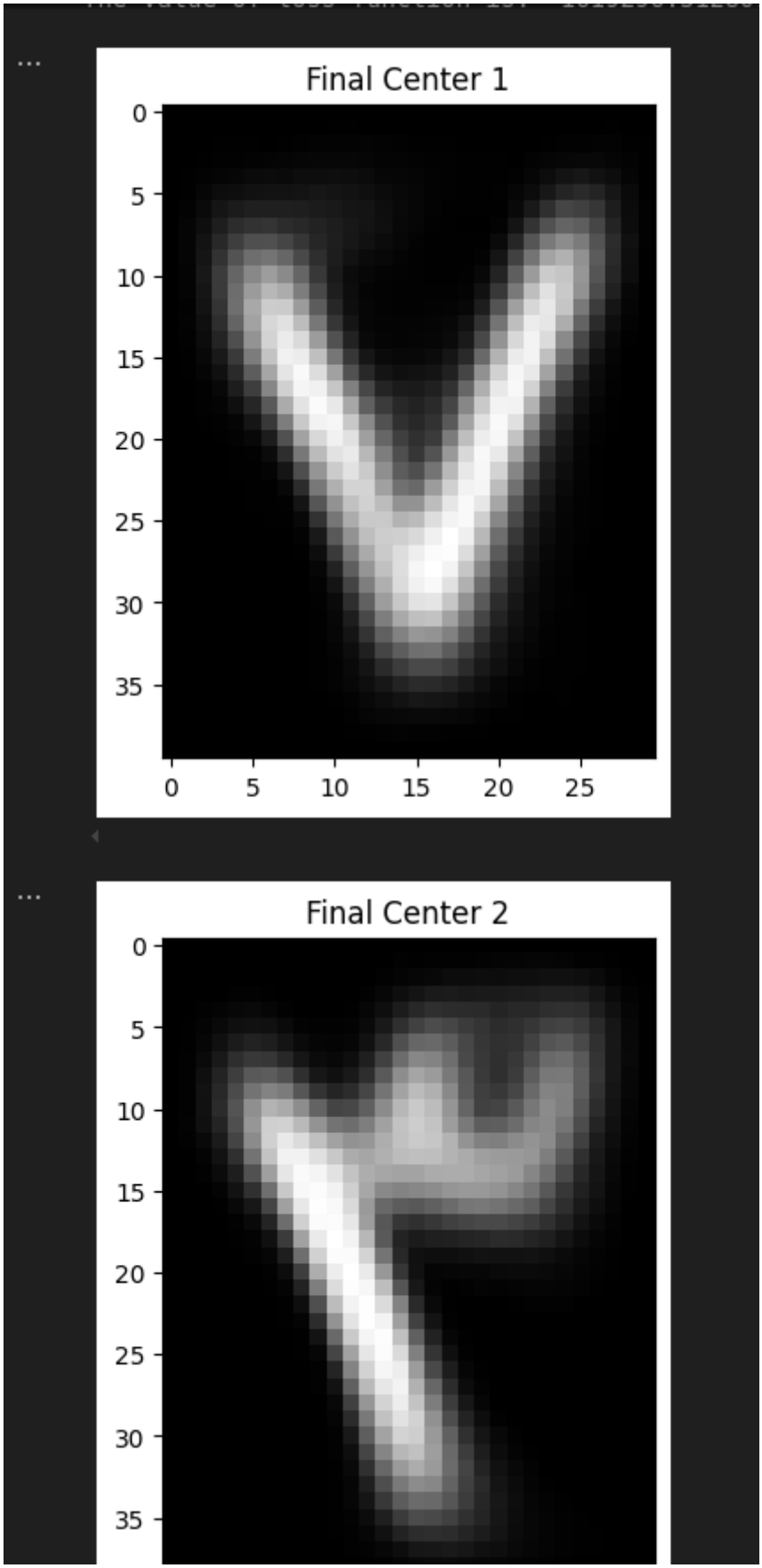

Initial Center 15
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial Center 16
0
5
10
15
20
25
30
35

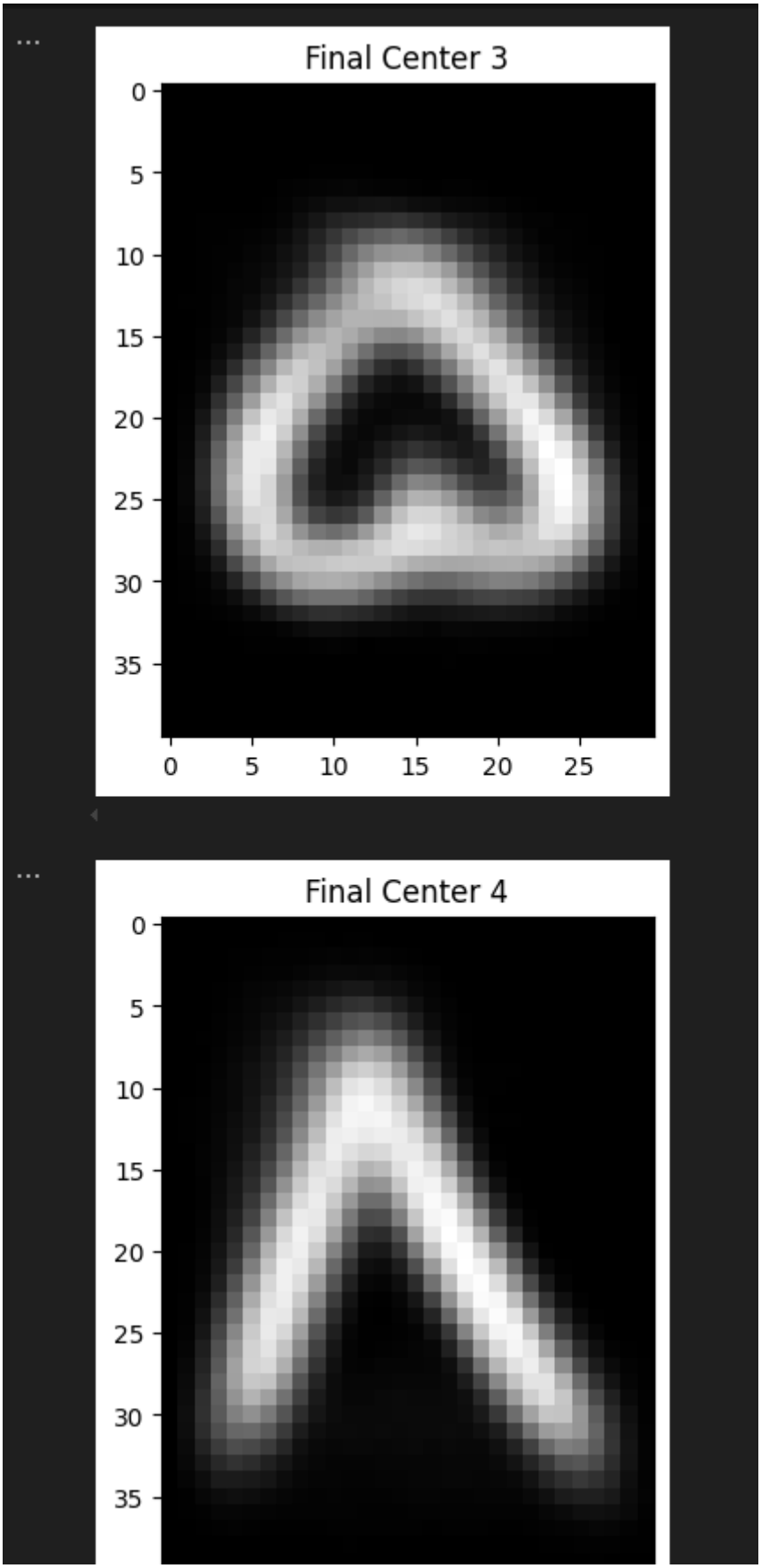

Initial Center 17
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial Center 18
0
5
10
15
20
25
30
35

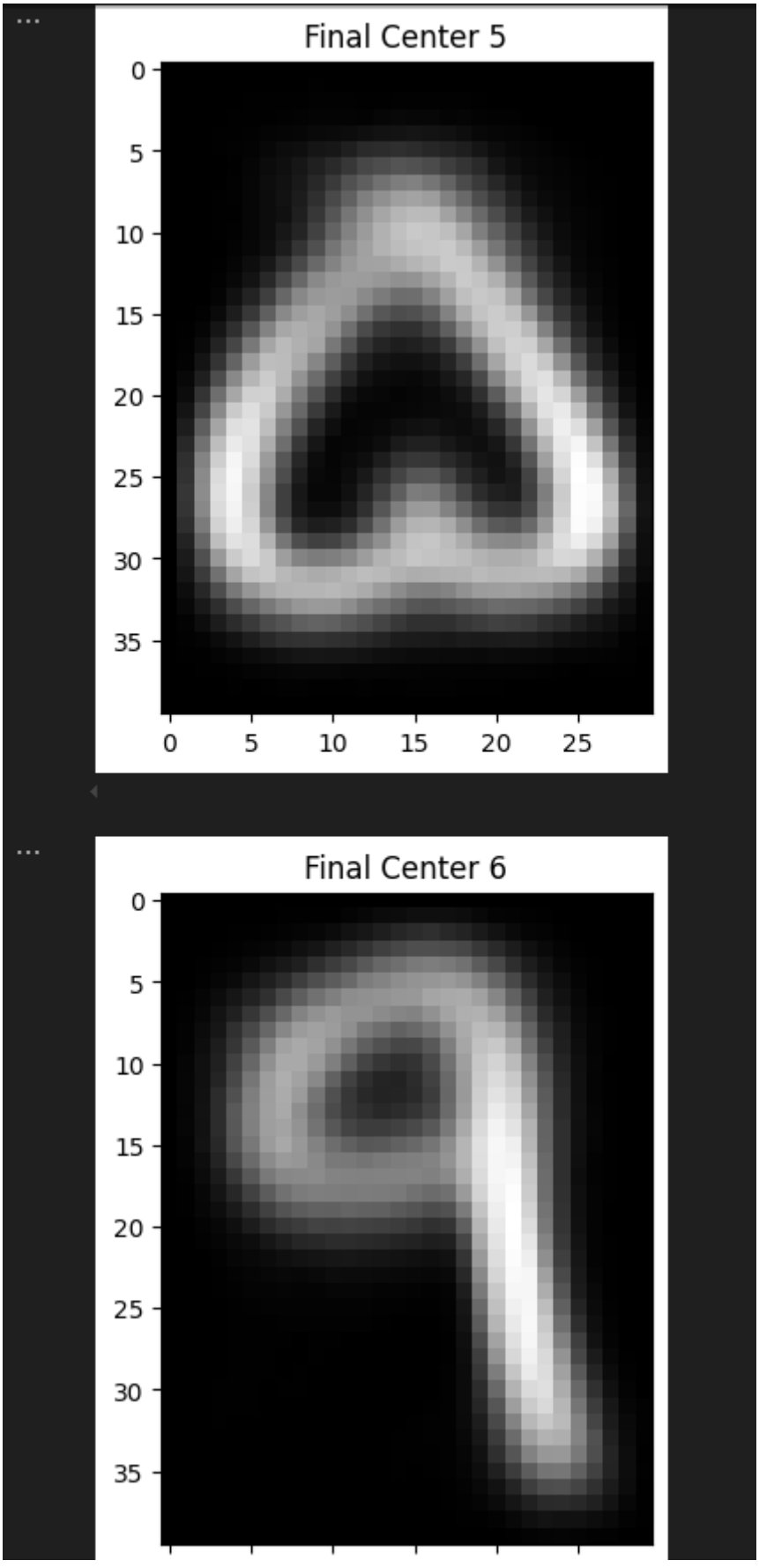

Initial Center 19
0
5
10
15
20
25
30
35
0
5
10
15
20
25

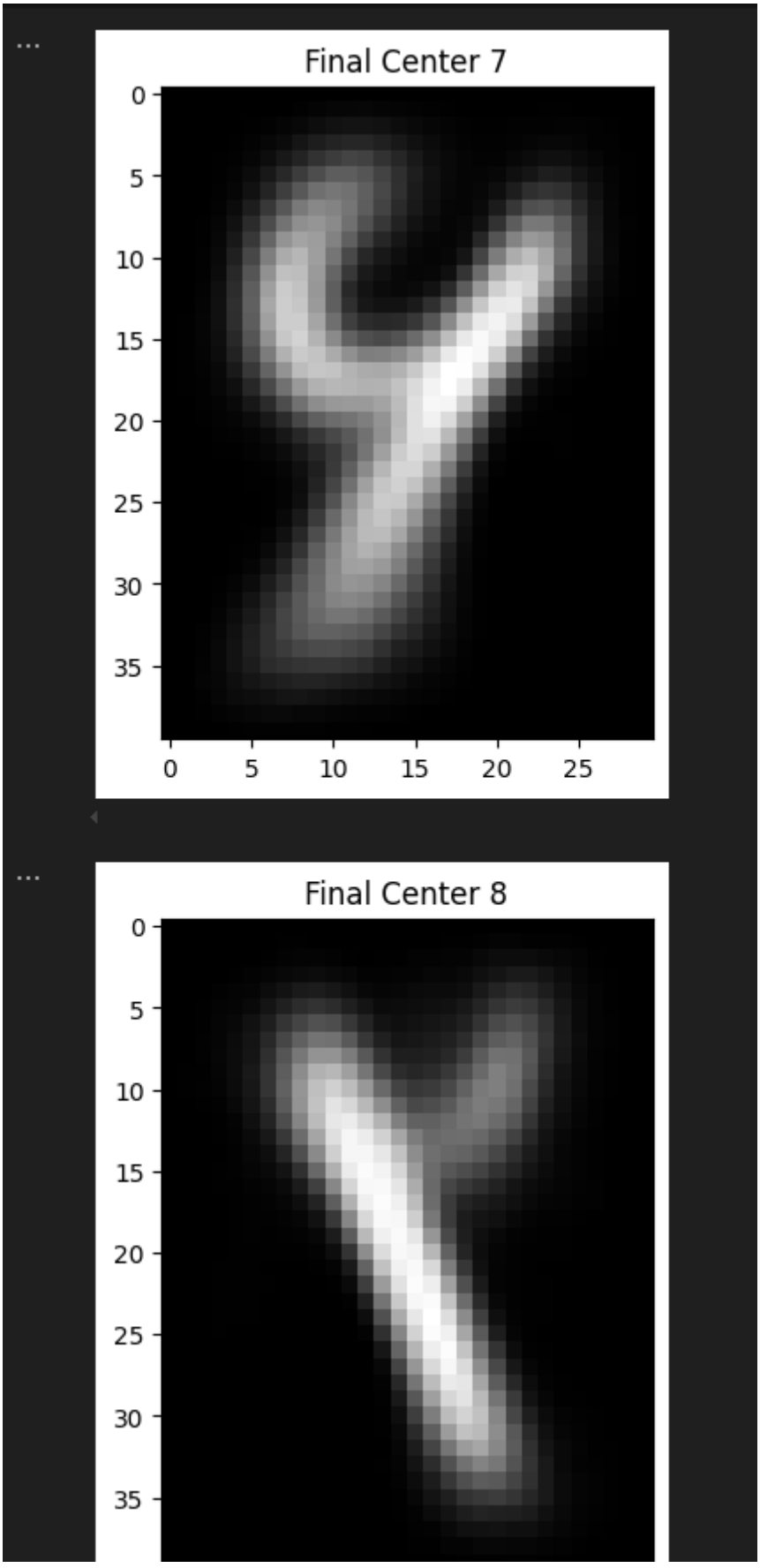

Initial Center 20
0
5
10
15
20
25
30
35

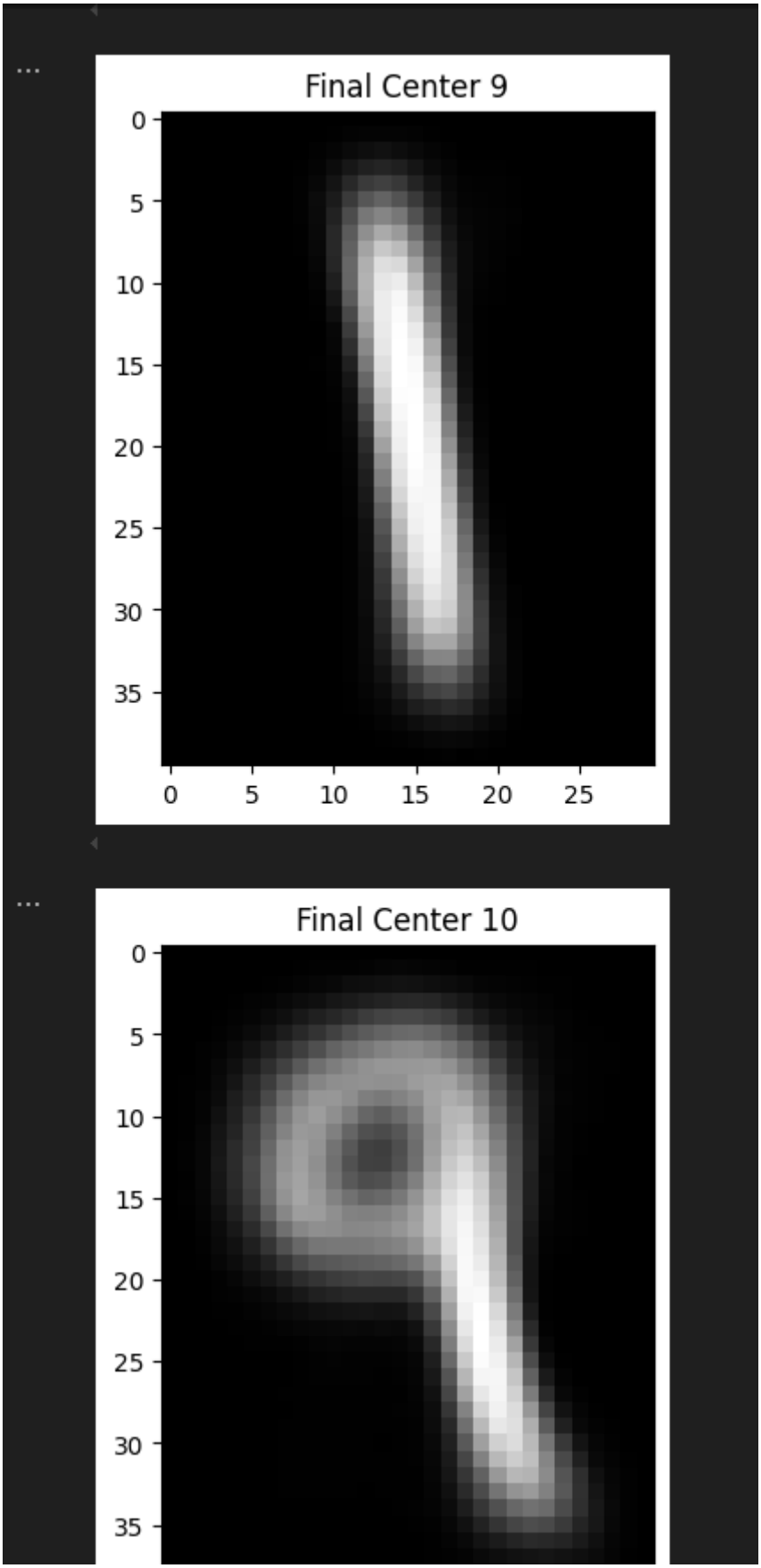
شاخص های نهایی

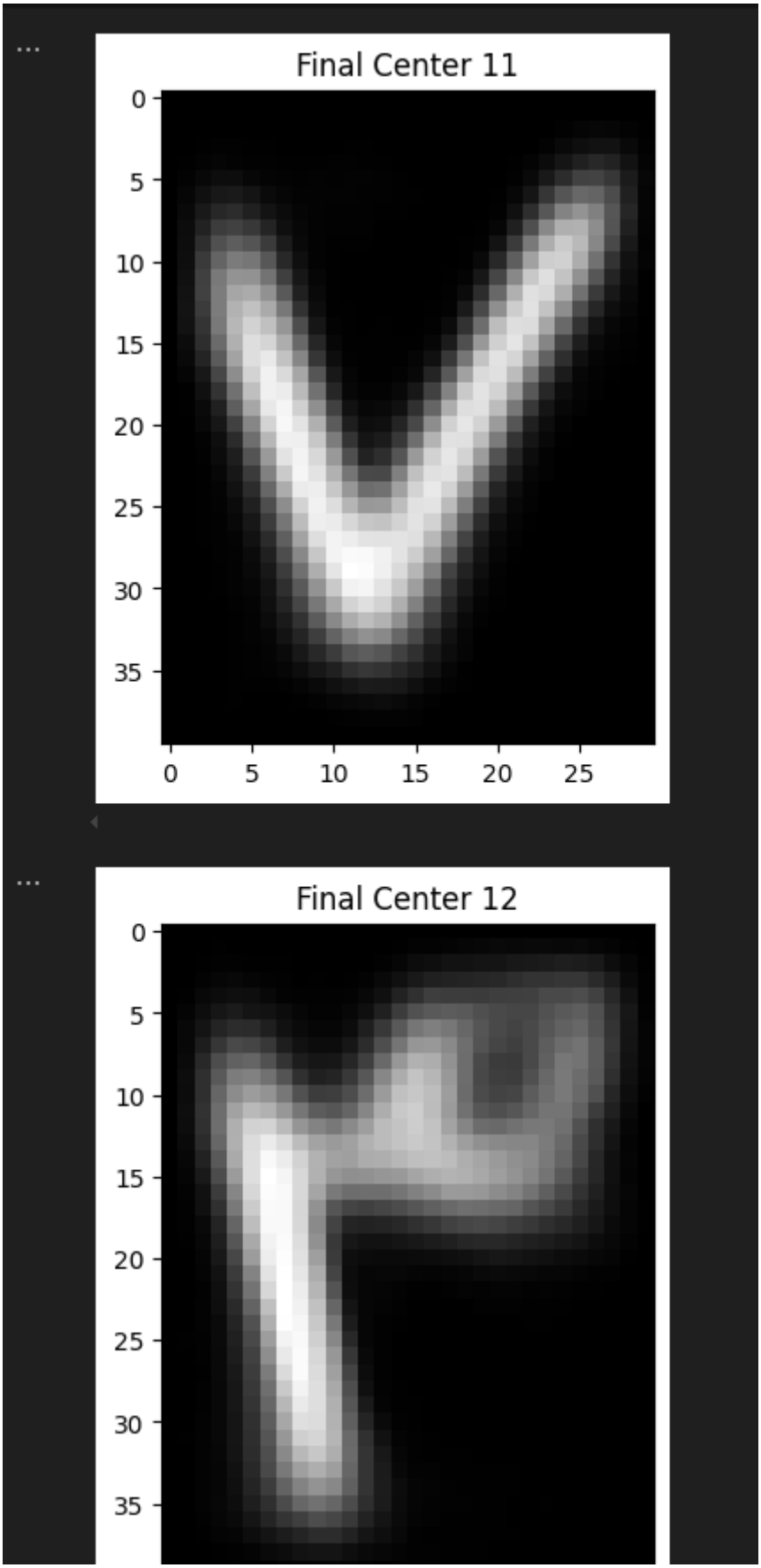


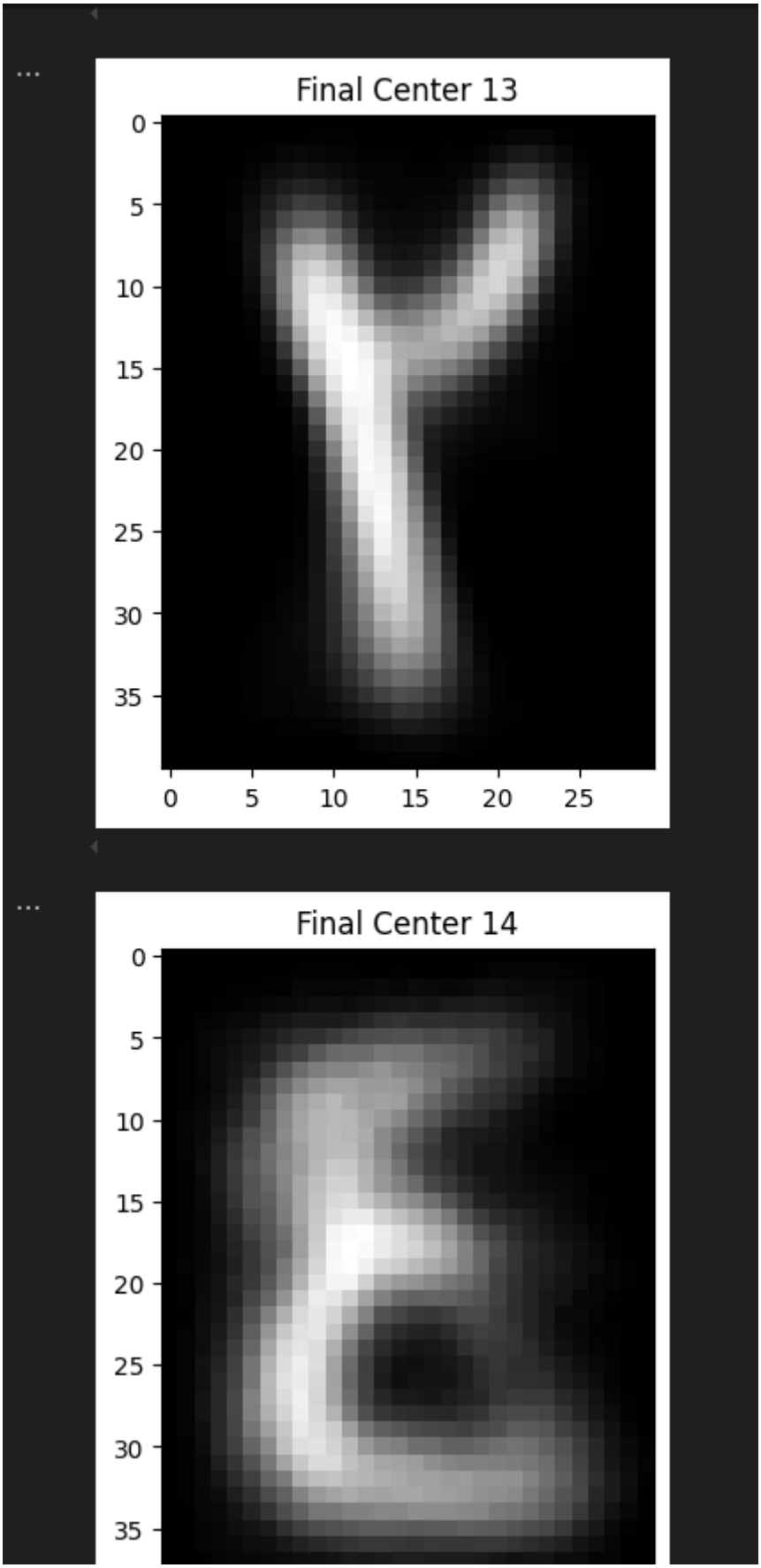

Final Center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 4
0
5
10
15
20
25
30
35

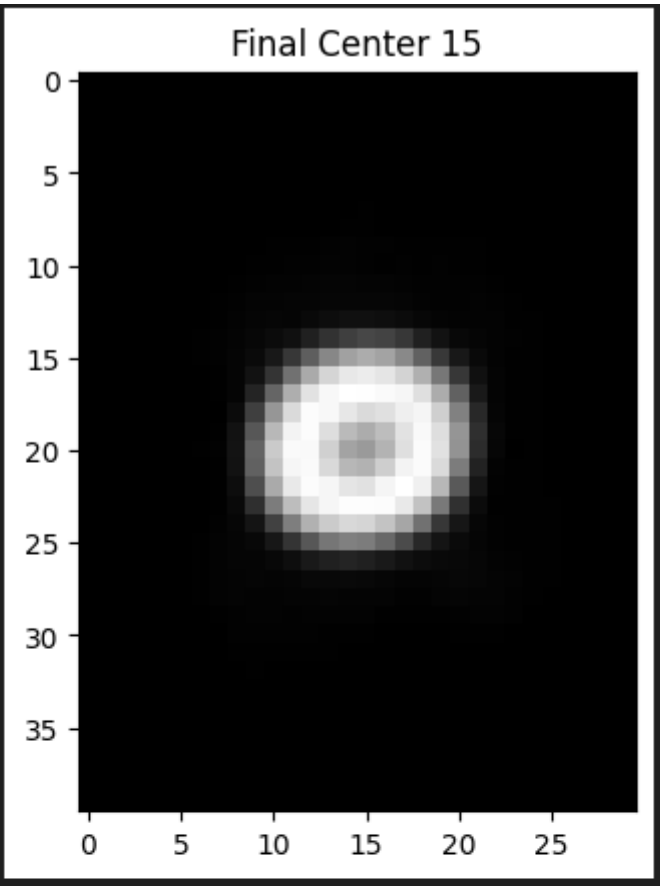

Final Center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 6
0
5
10
15
20
25
30
35

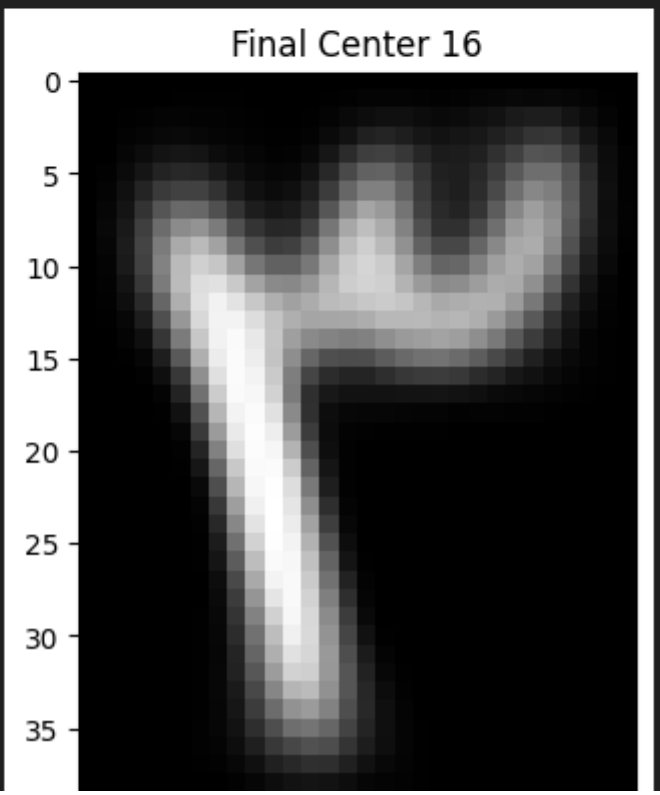

Final Center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 8
0
5
10
15
20
25
30
35

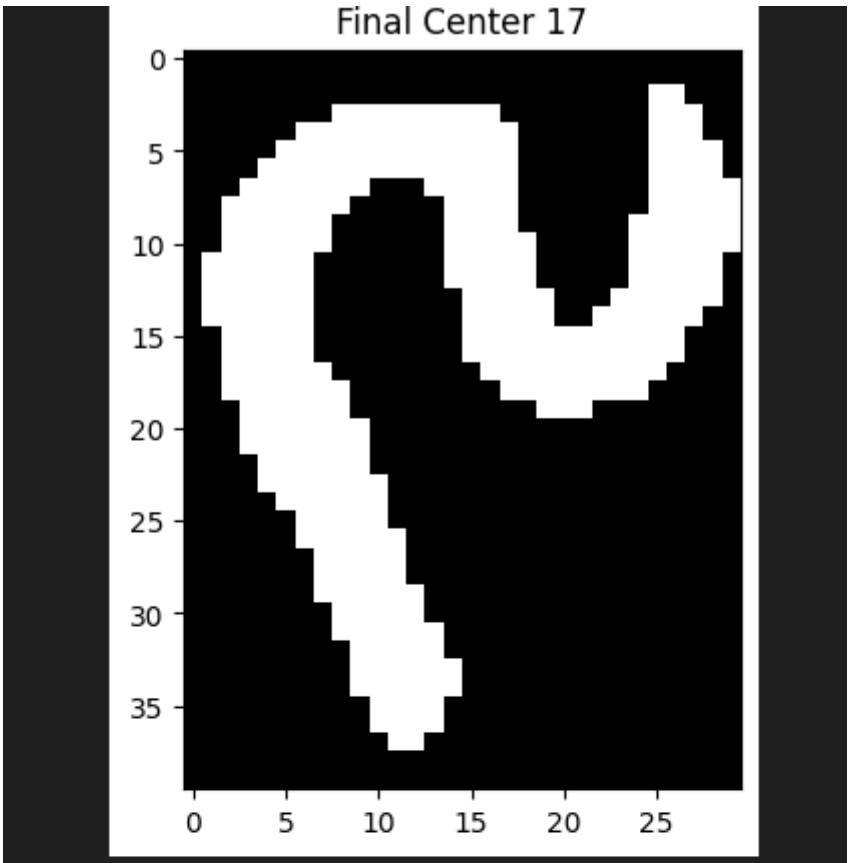

Final Center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 10
0
5
10
15
20
25
30
35


Final Center 11
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 12
0
5
10
15
20
25
30
35

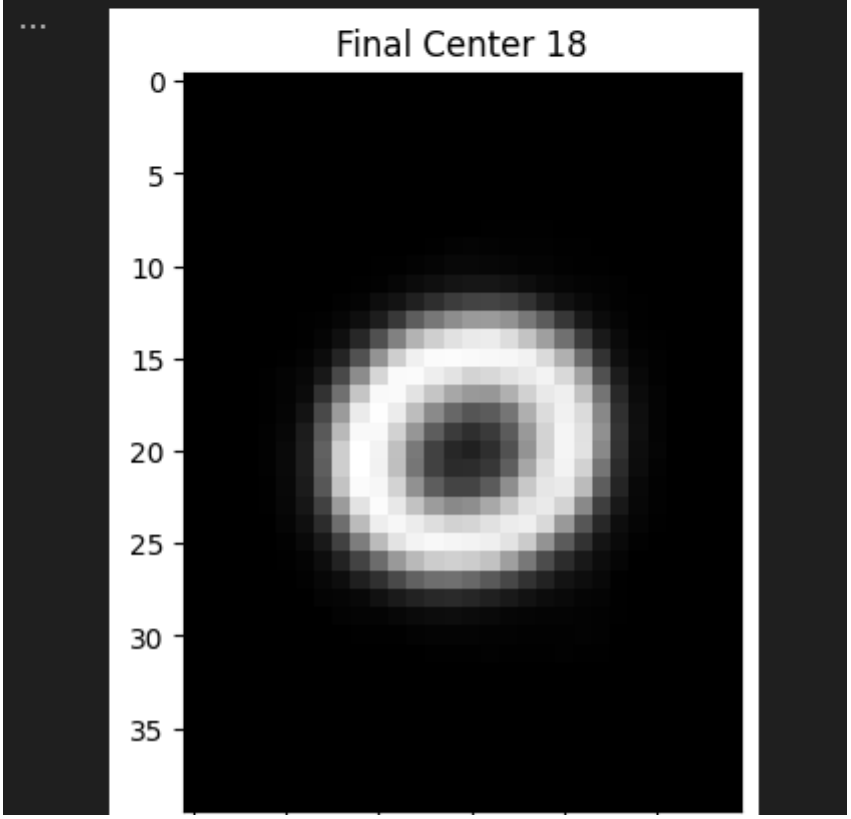

Final Center 13
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 14
0
5
10
15
20
25
30
35

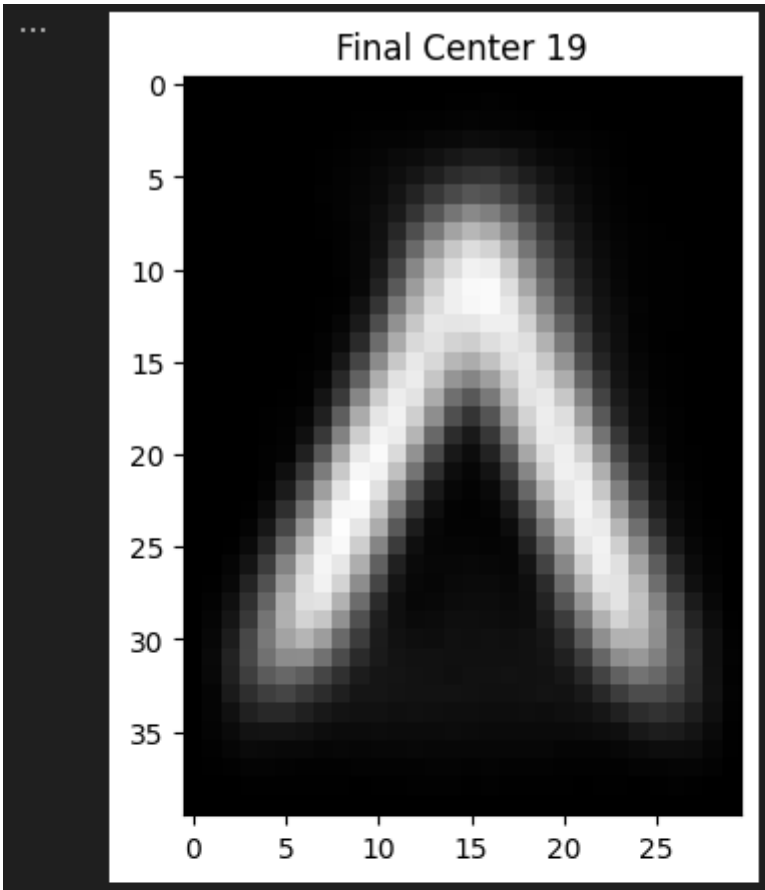

Final Center 15
0
5
10
15
20
25
30
35
0
5
10
15
20
25

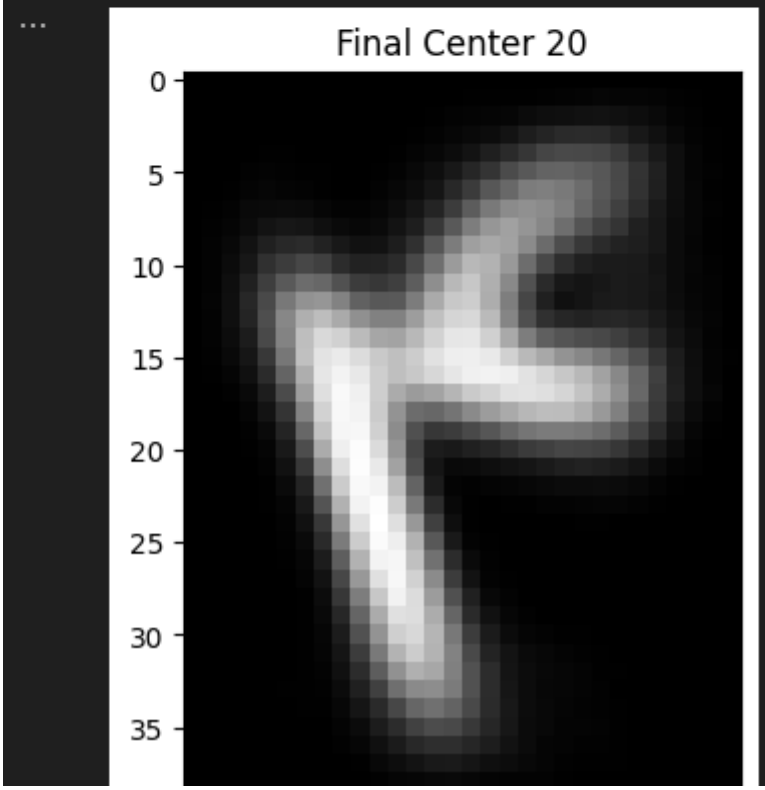

Final Center 16
0
5
10
15
20
25
30
35


Final Center 17
0
5
10
15
20
25
30
35
0
5
10
15
20
25

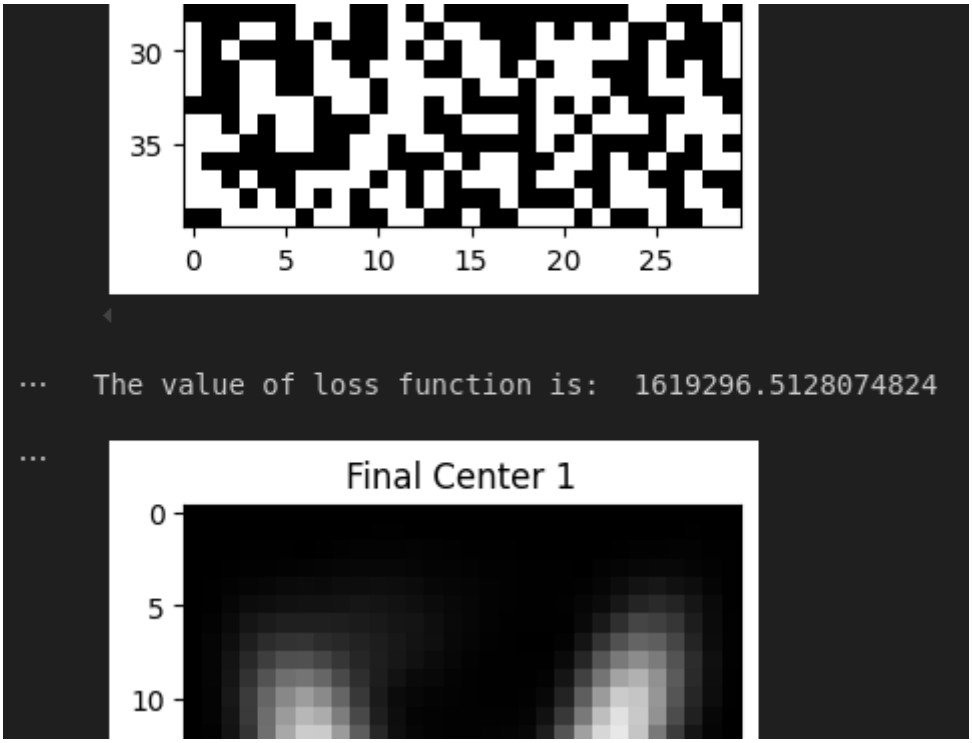
...


Final Center 18
0
5
10
15
20
25
30
35


Final Center 19
0
5
10
15
20
25
30
35
0
5
10
15
20
25

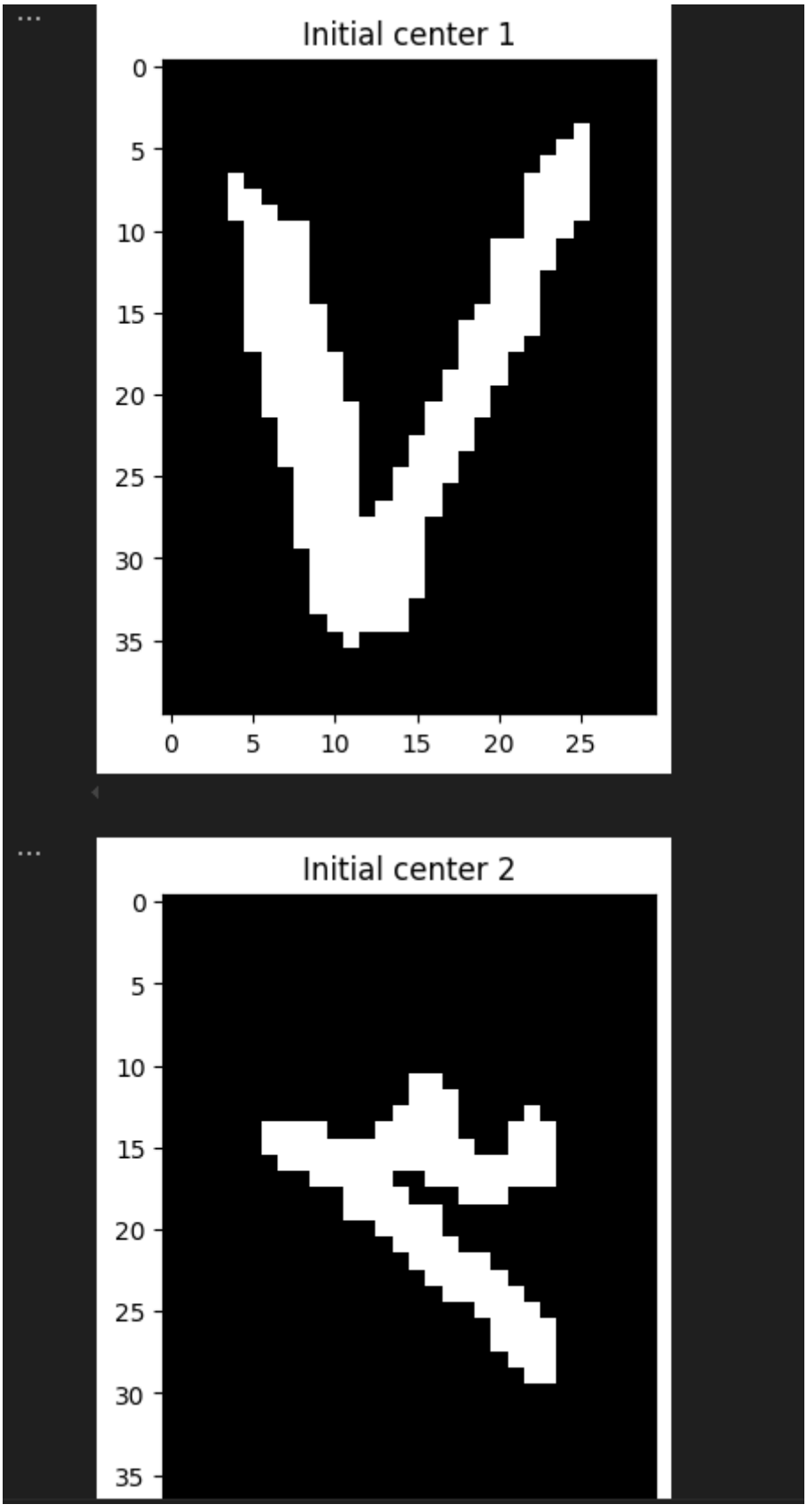

Final Center 20
0
5
10
15
20
25
30
35

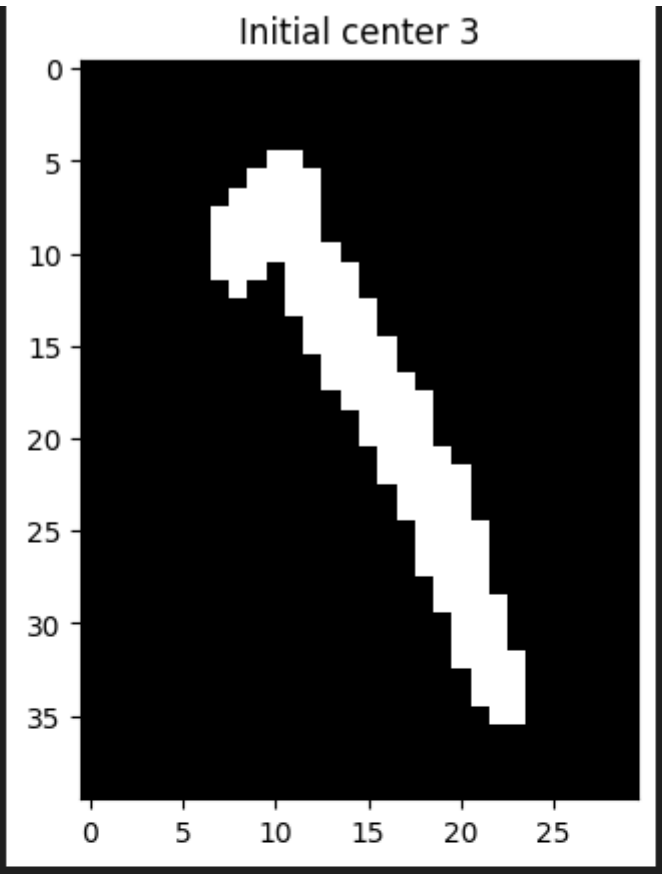
مقدار تابع هزینه

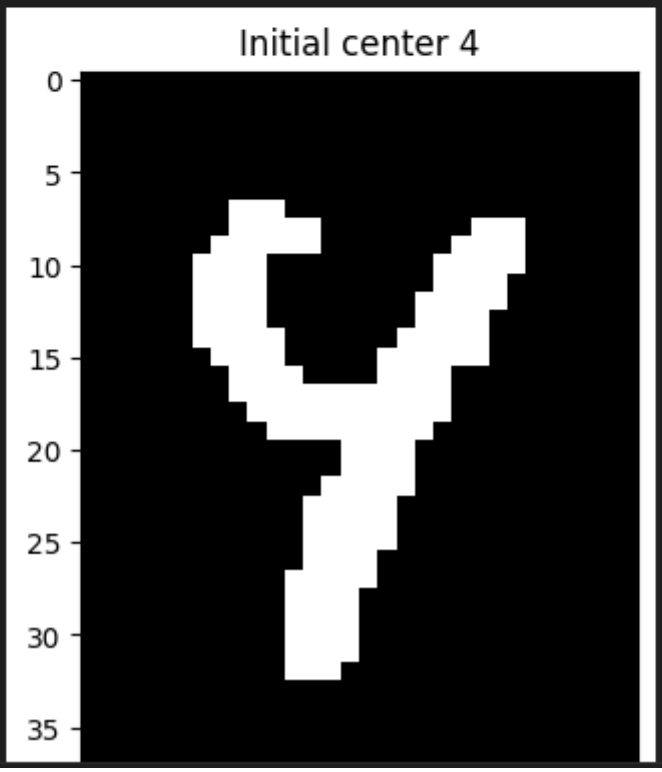


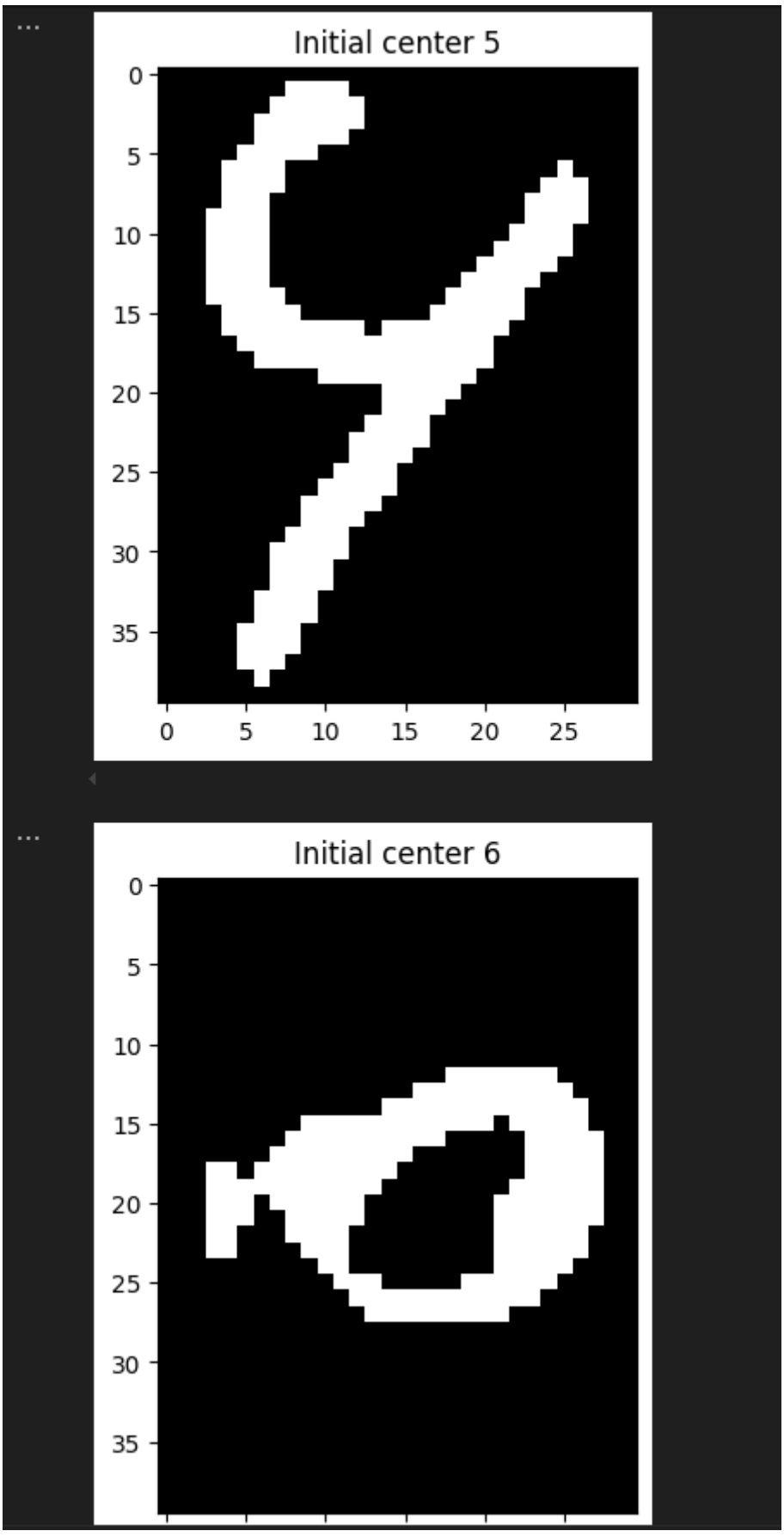
b.II:

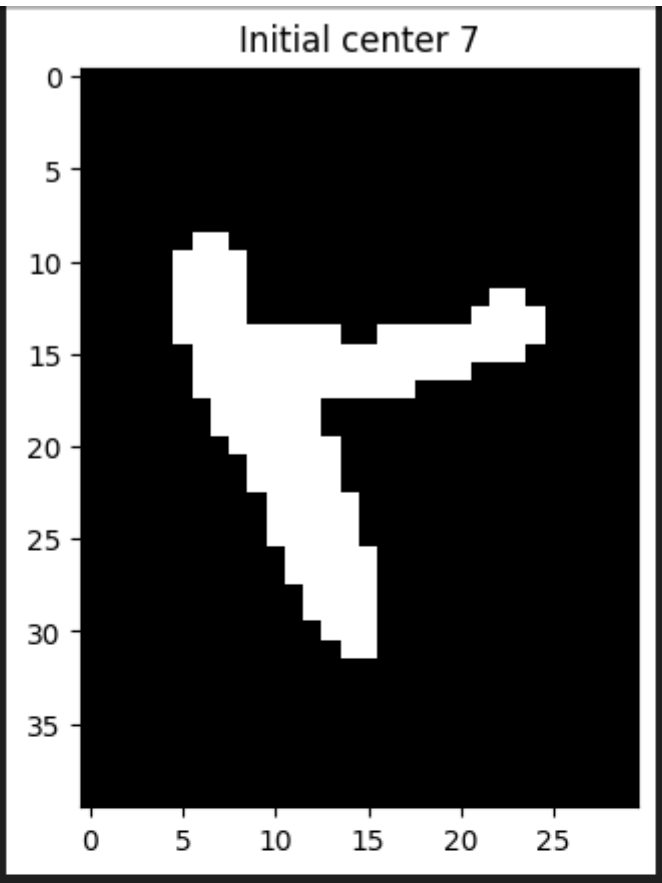
شاخص های اولیه تصادفی انتخاب شده از بین داده ها

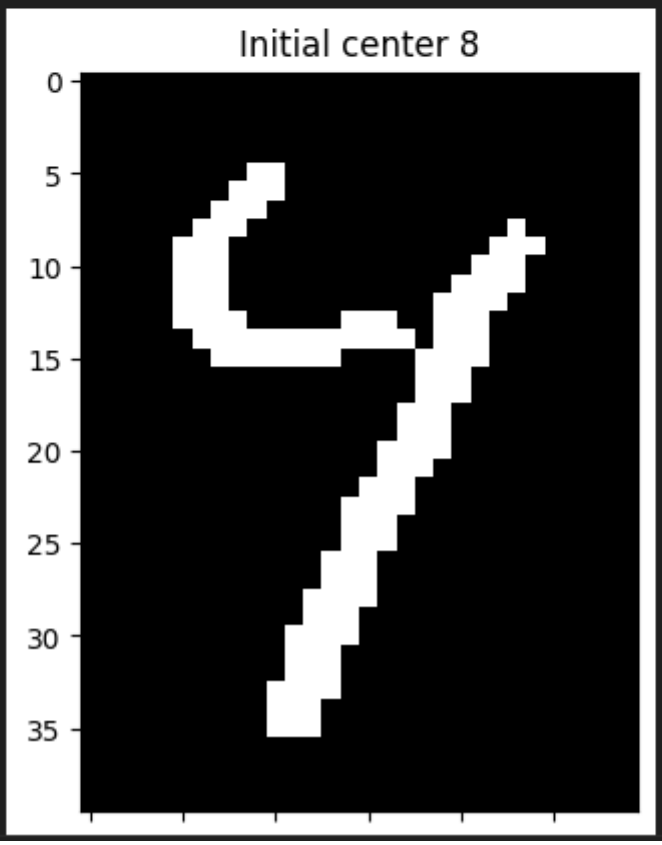


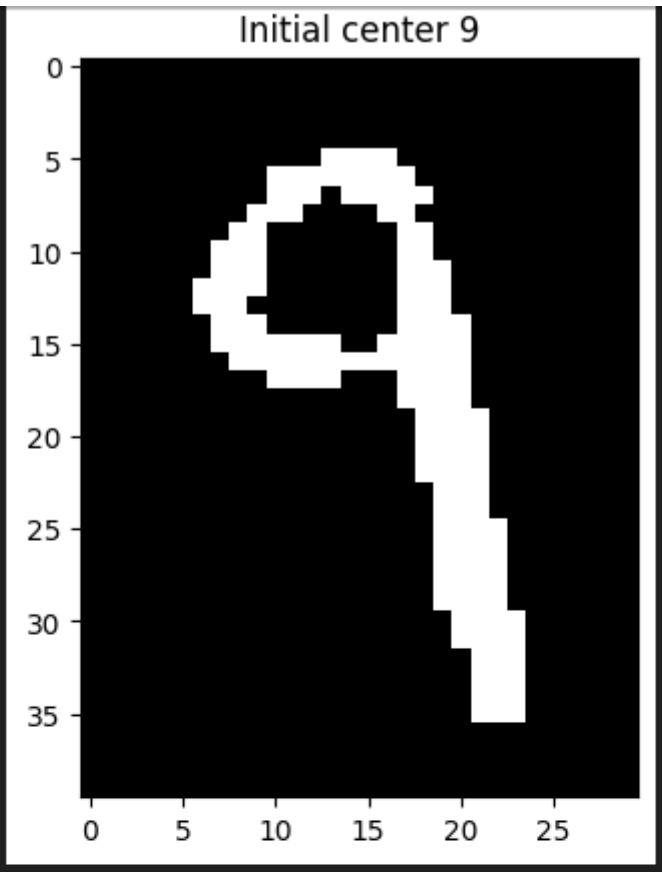

Initial center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25

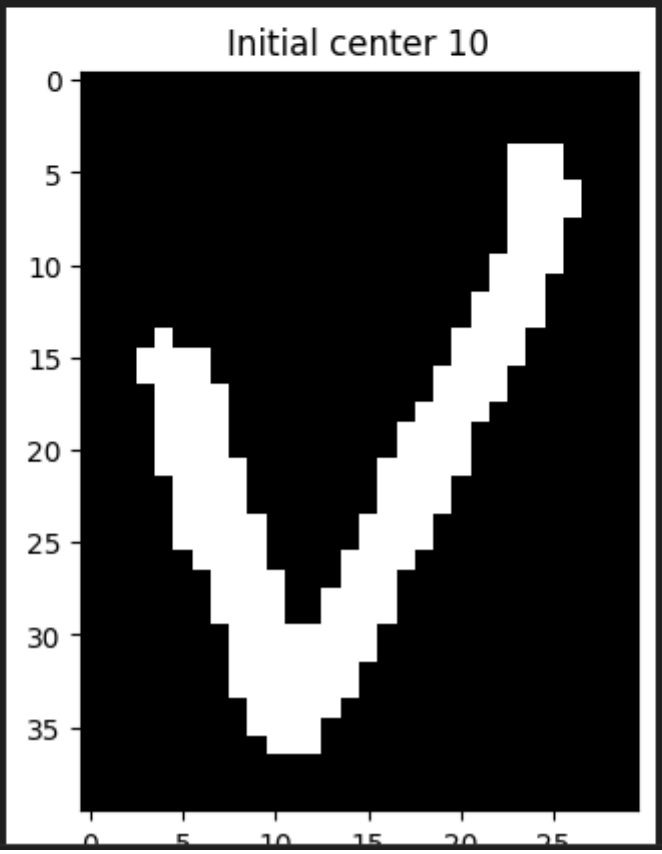

Initial center 4
0
5
10
15
20
25
30
35

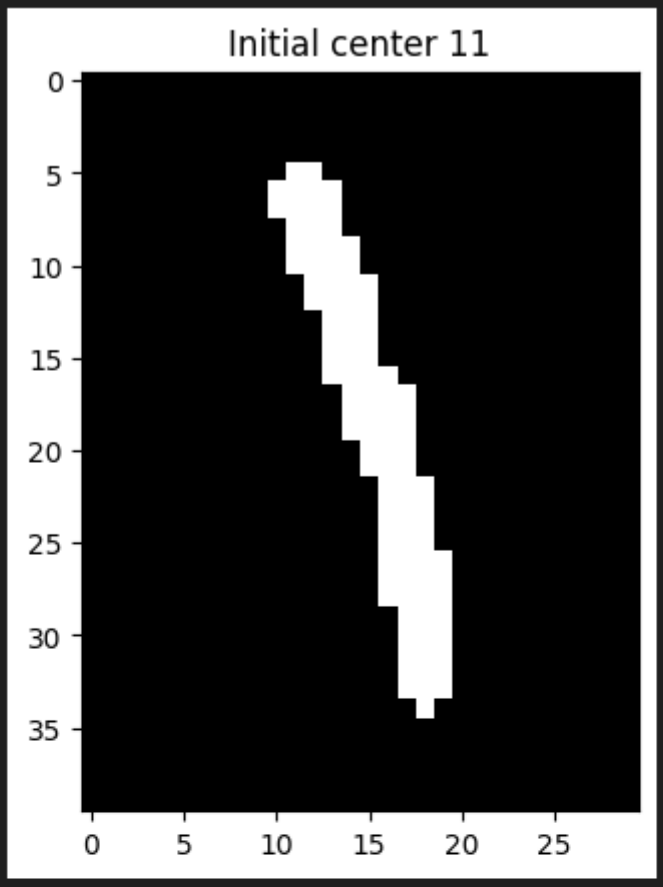

Initial center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial center 6
0
5
10
15
20
25
30
35

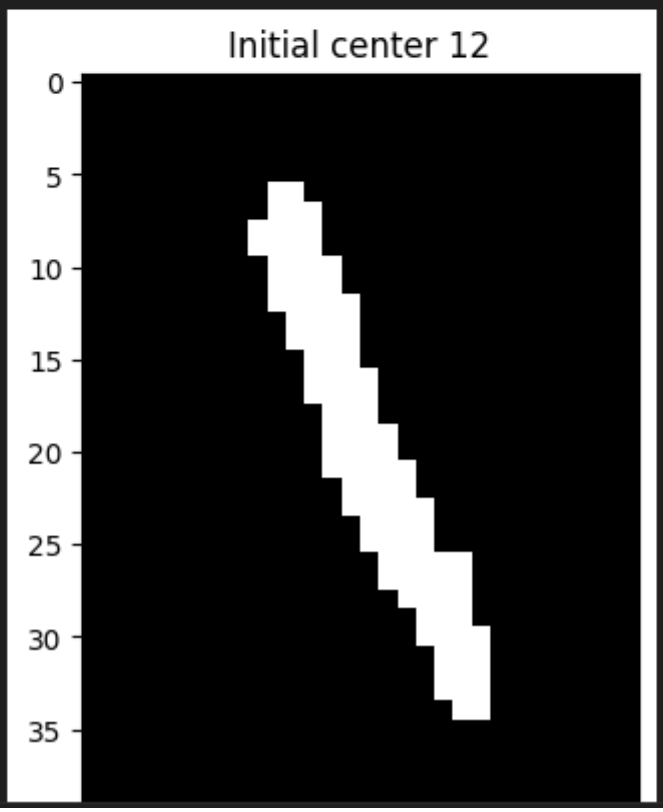

Initial center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25

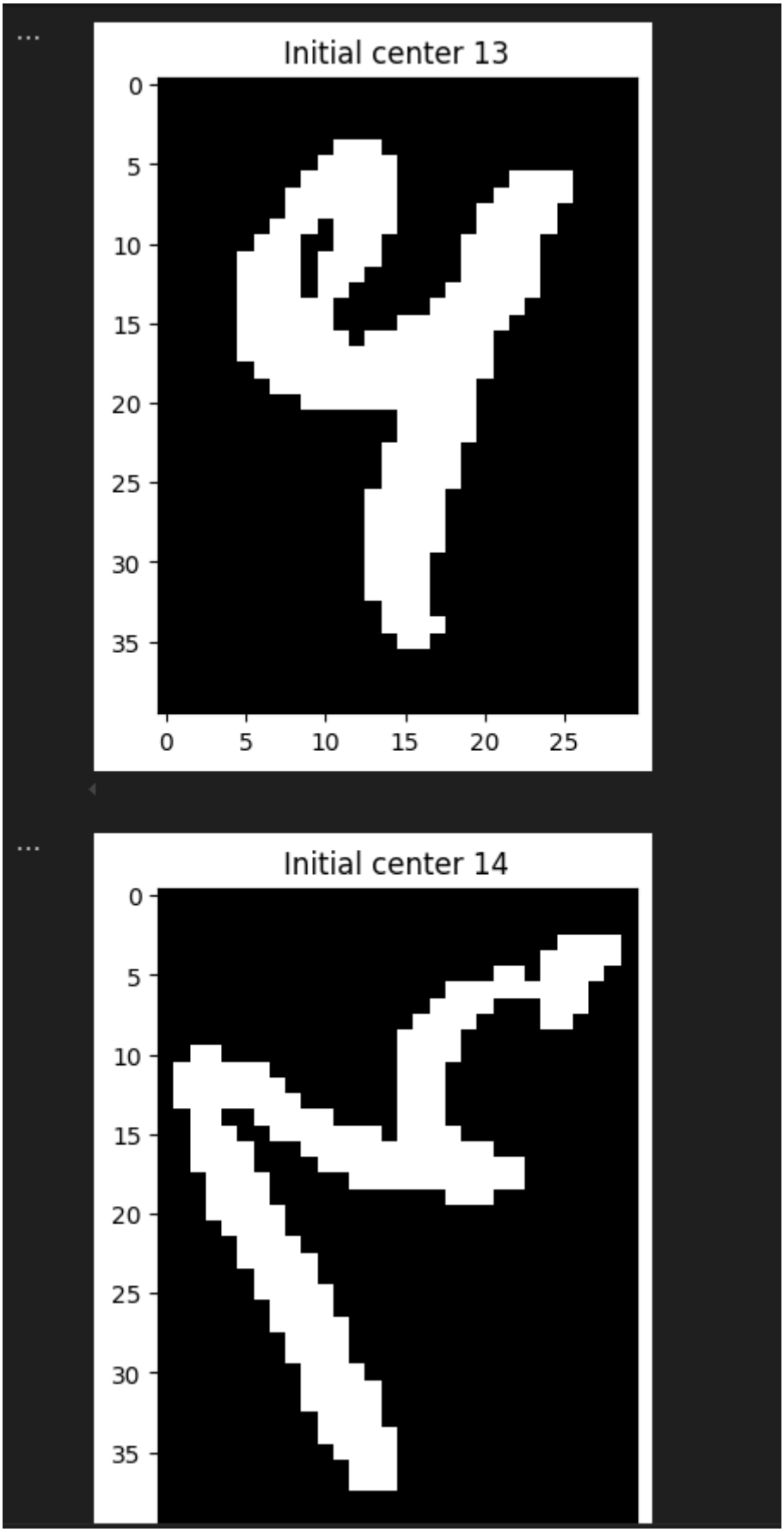

Initial center 8
0
5
10
15
20
25
30
35


Initial center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Initial center 10
0
5
10
15
20
25
30
35

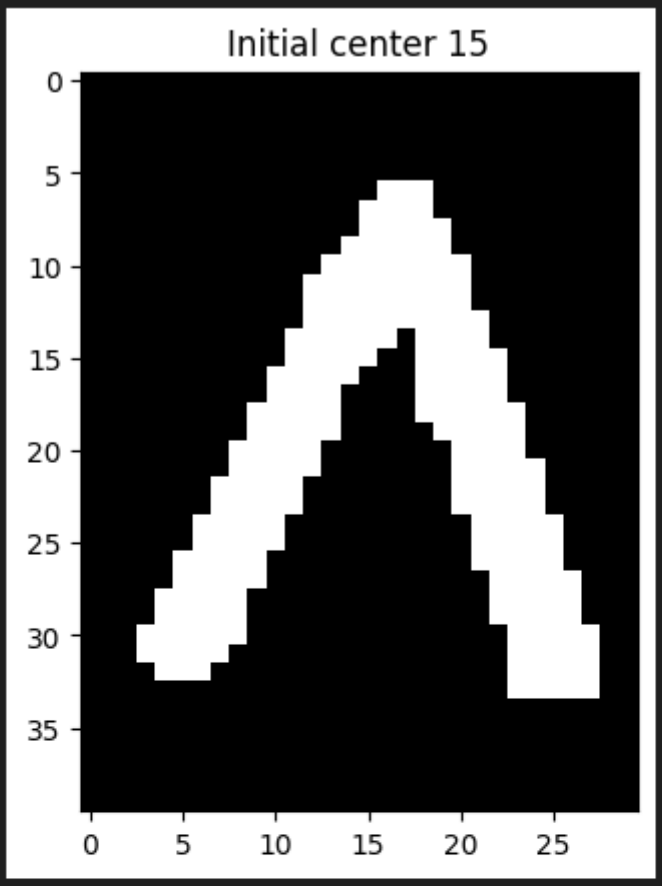

Initial center 11
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Initial center 12
0
5
10
15
20
25
30
35

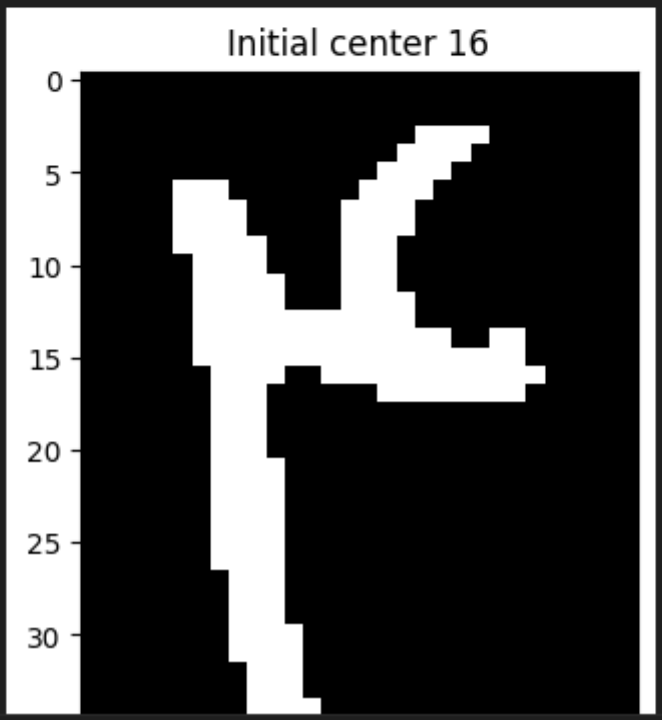

Initial center 13
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial center 14
0
5
10
15
20
25
30
35

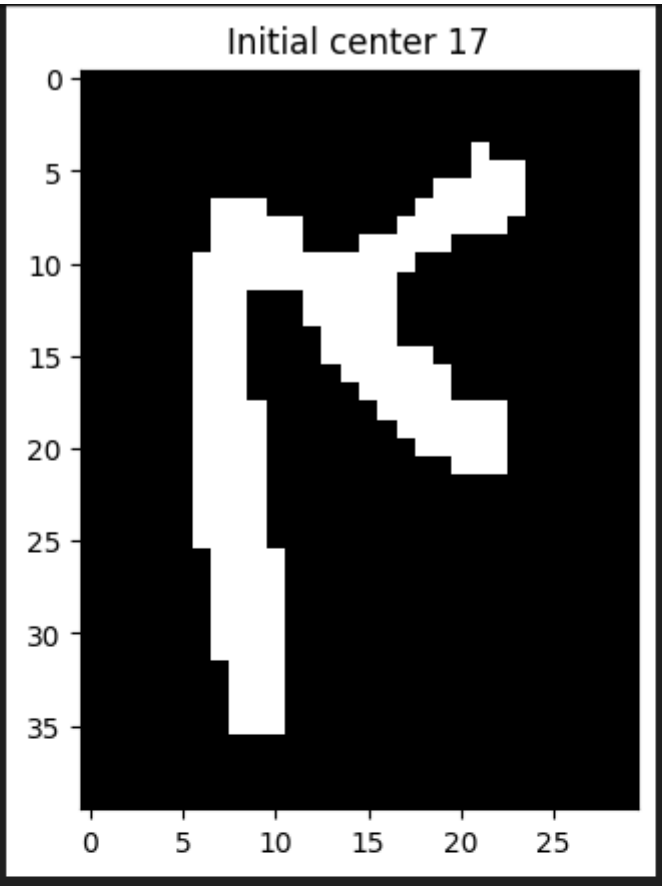

0
5
10
15
20
25

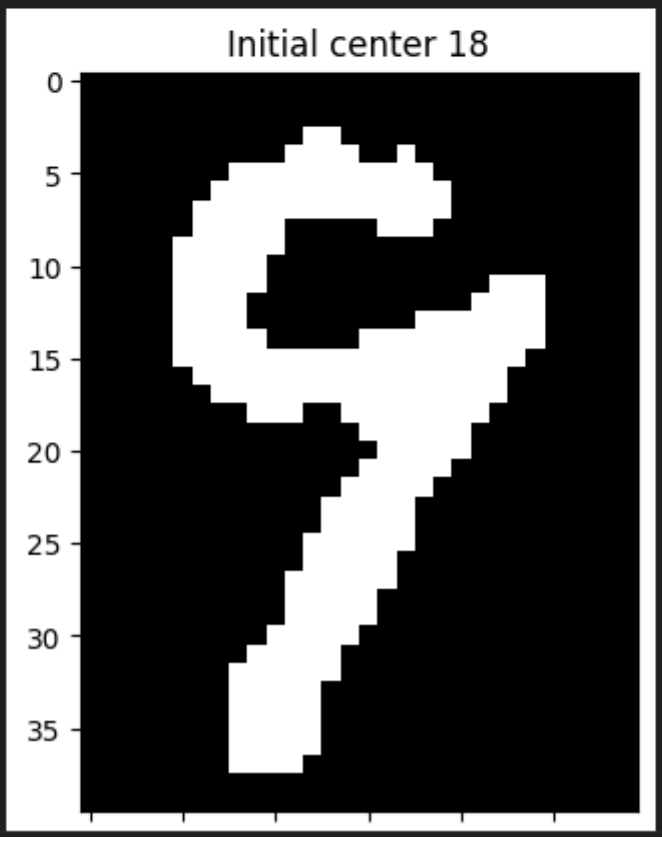
...

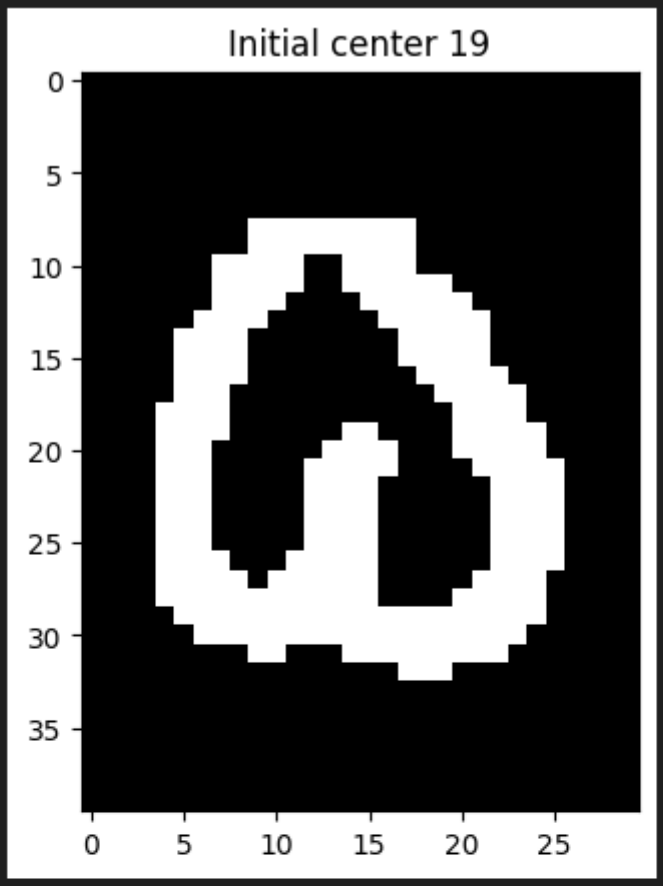

Initial center 15
0
5
10
15
20
25
30
35
0
5
10
15
20
25

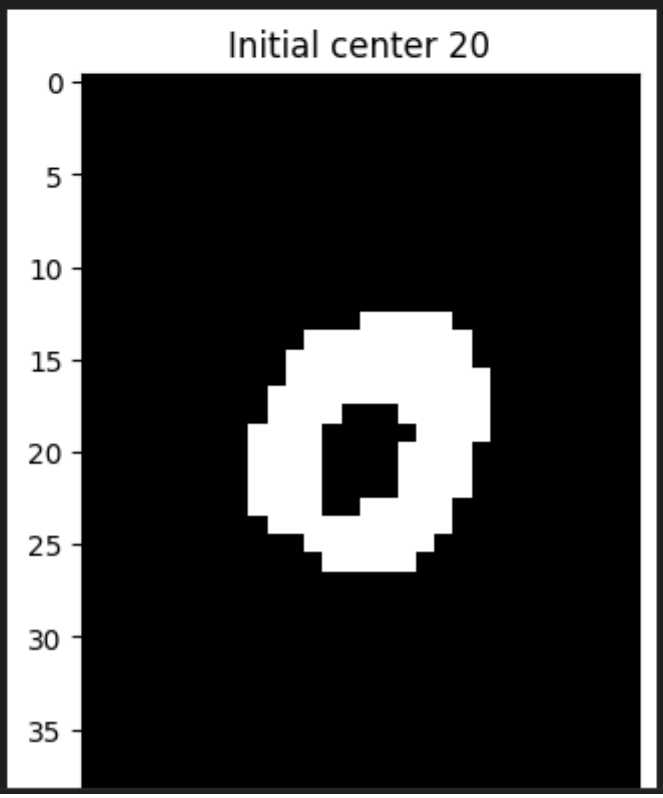
...


Initial center 16
0
5
10
15
20
25
30

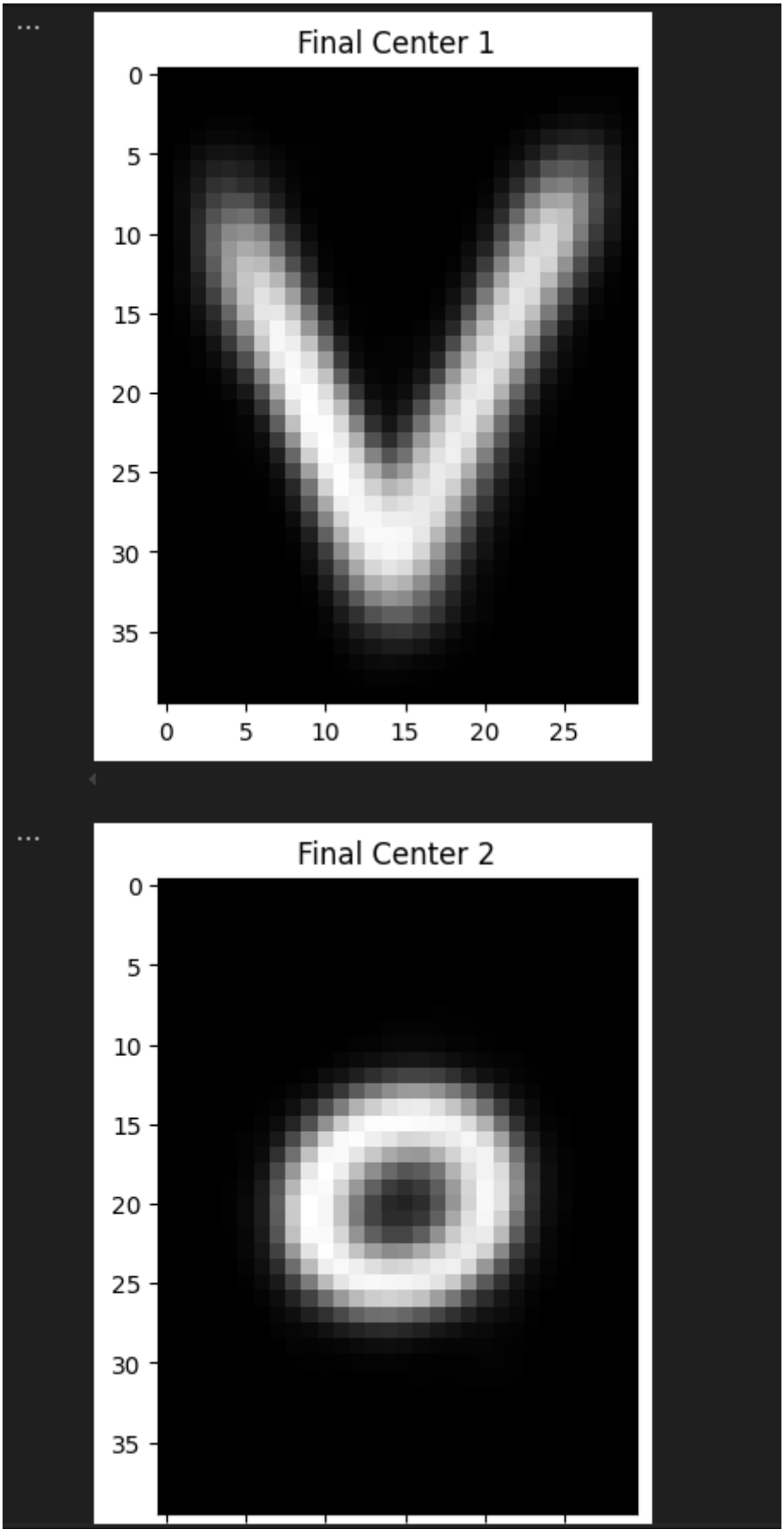

Initial center 17
0
5
10
15
20
25
30
35
0
5
10
15
20
25

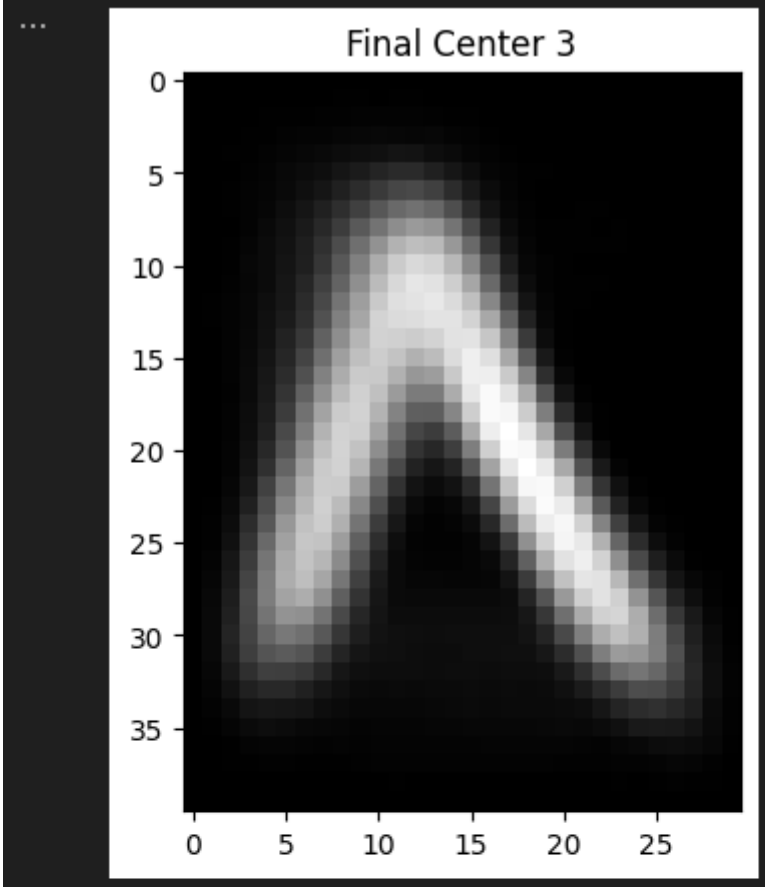

Initial center 18
0
5
10
15
20
25
30
35

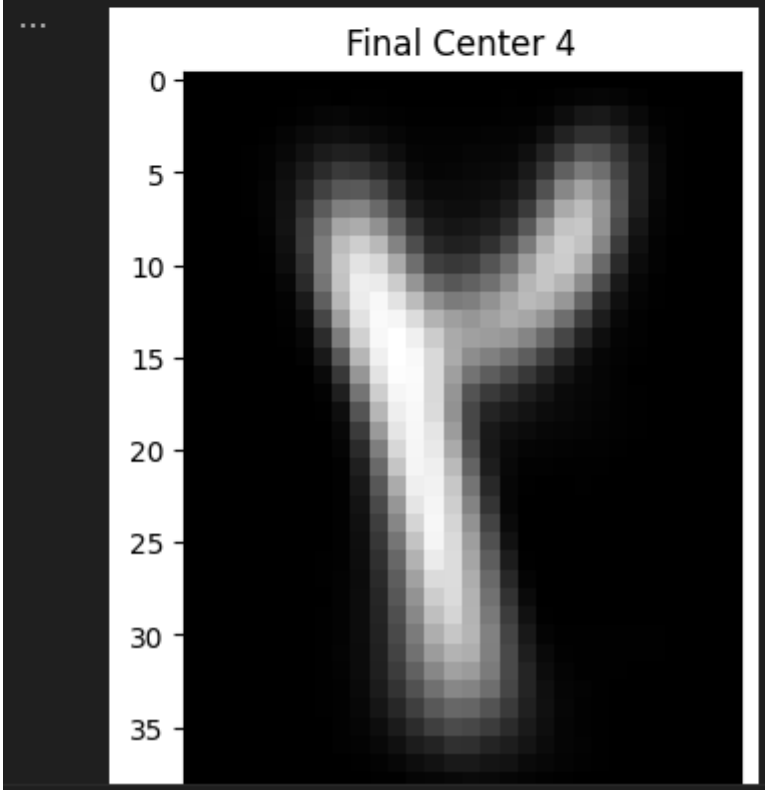

Initial center 19
0
5
10
15
20
25
30
35
0
5
10
15
20
25

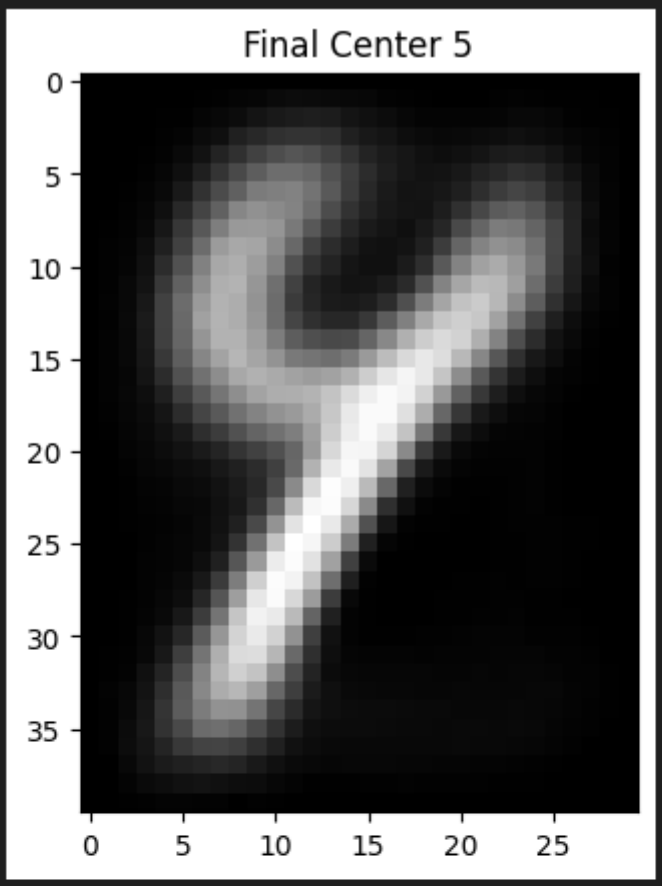

Initial center 20
0
5
10
15
20
25
30
35

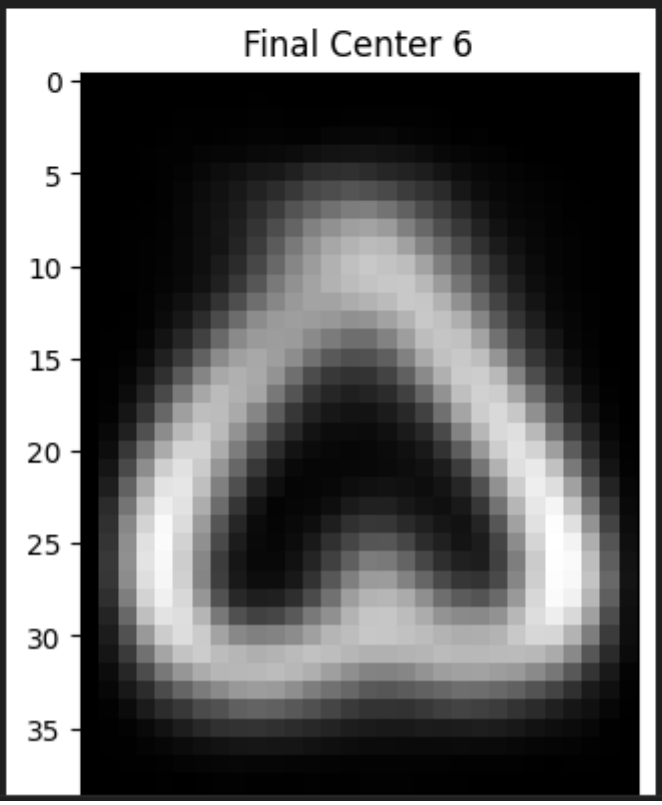
شاخص های نهایی

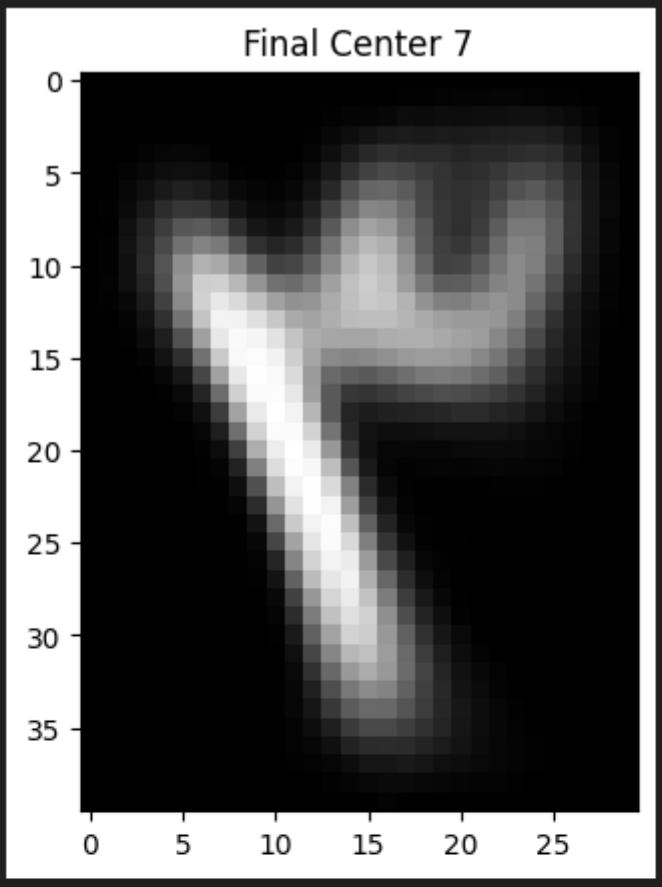


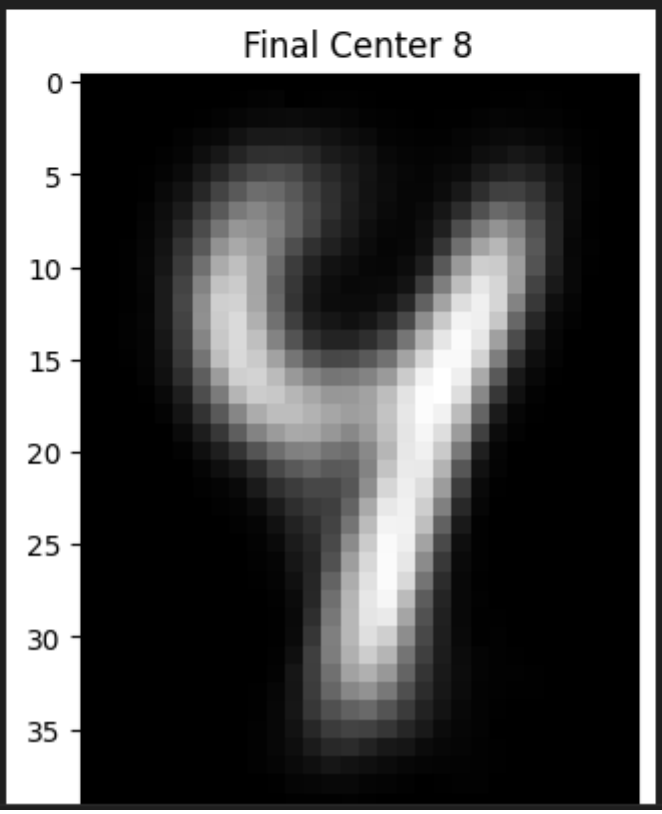

Final Center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25

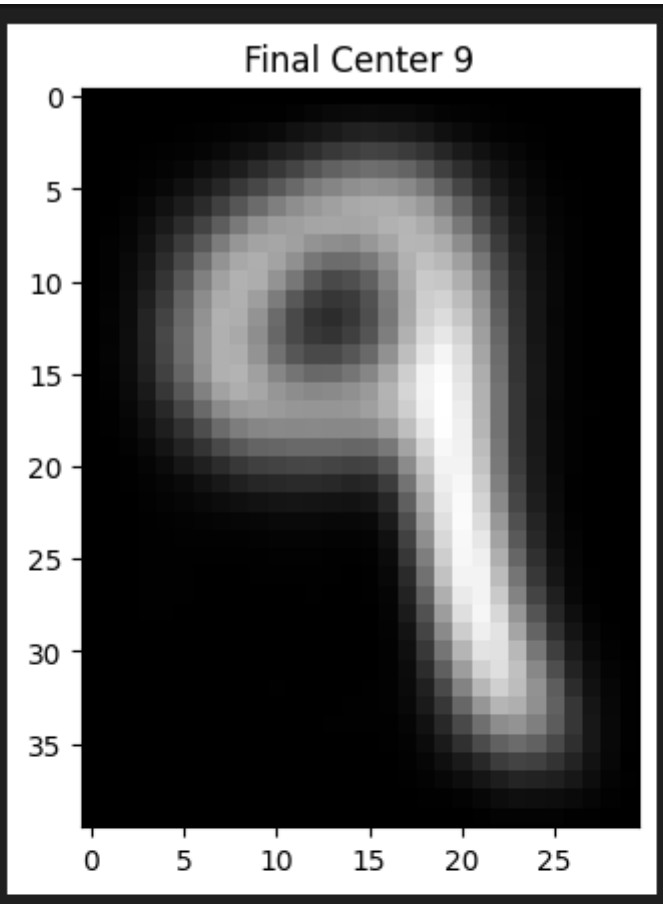

Final Center 4
0
5
10
15
20
25
30
35

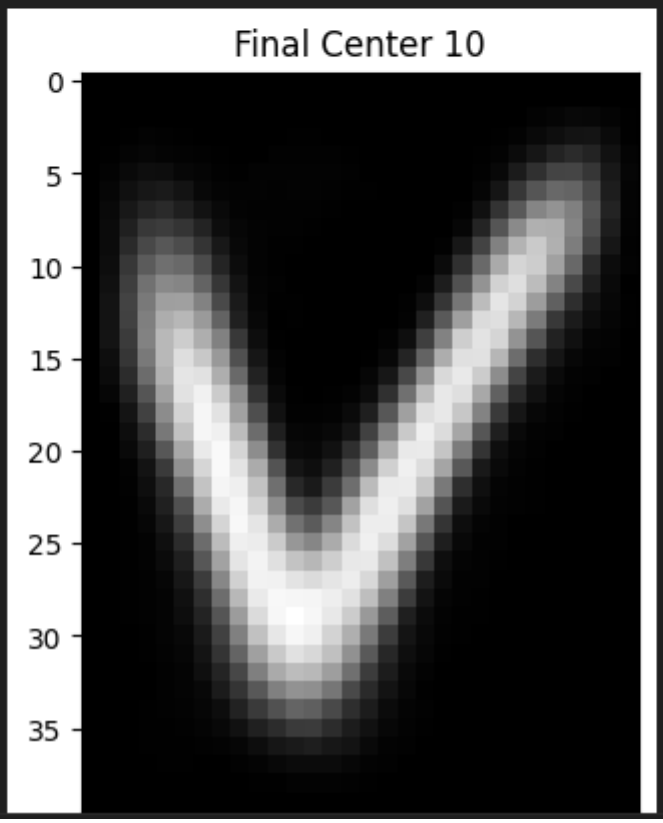

Final Center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25

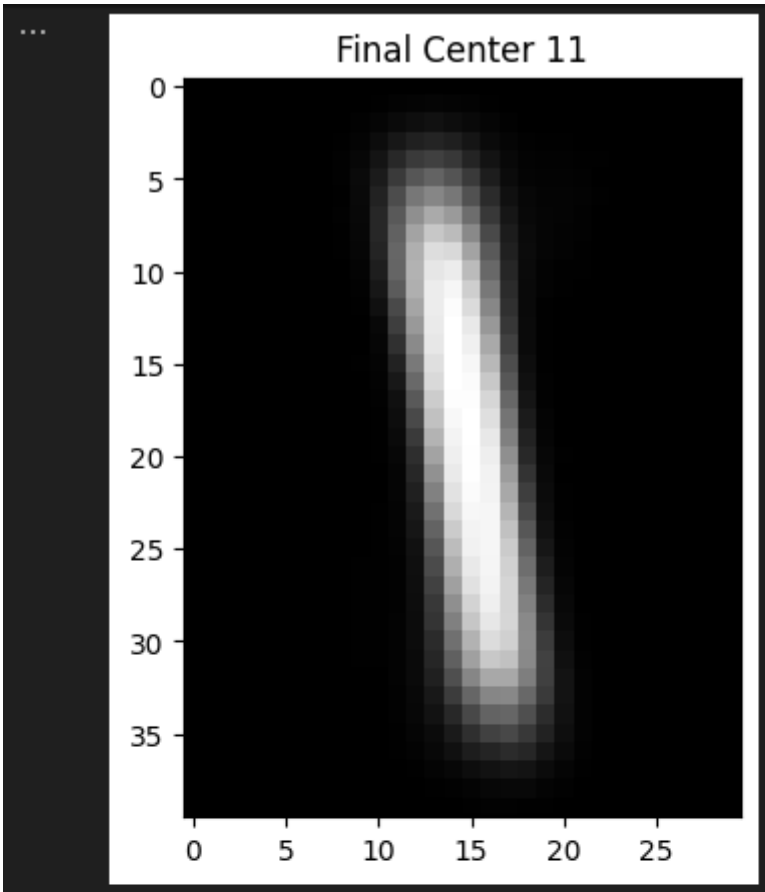

Final Center 6
0
5
10
15
20
25
30
35

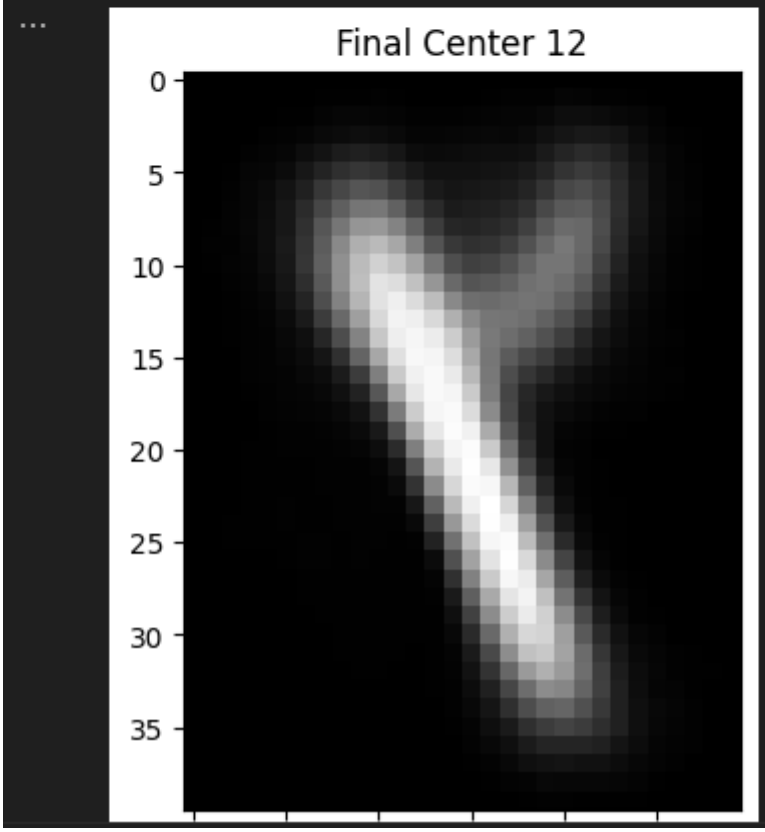

Final Center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25

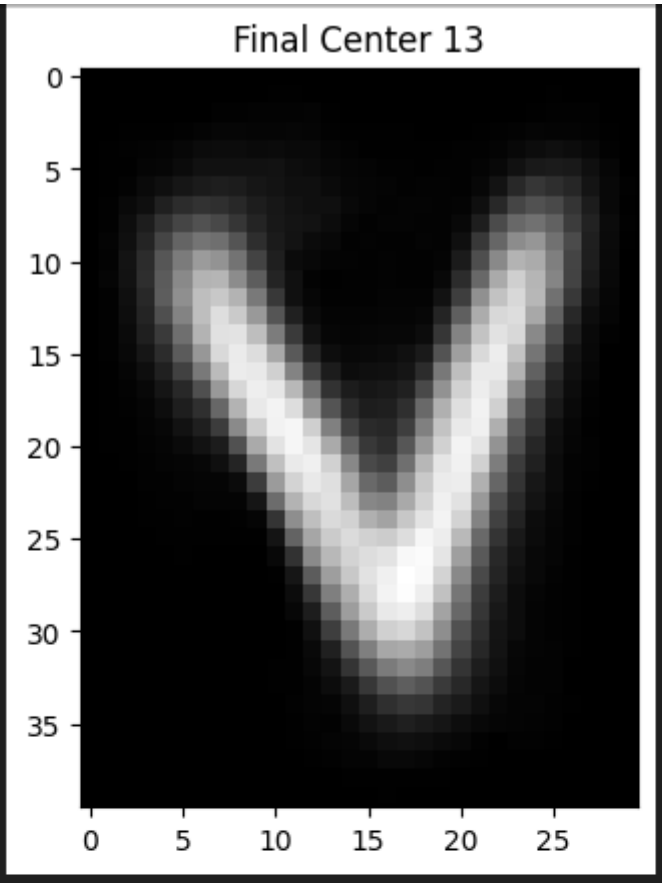

Final Center 8
0
5
10
15
20
25
30
35

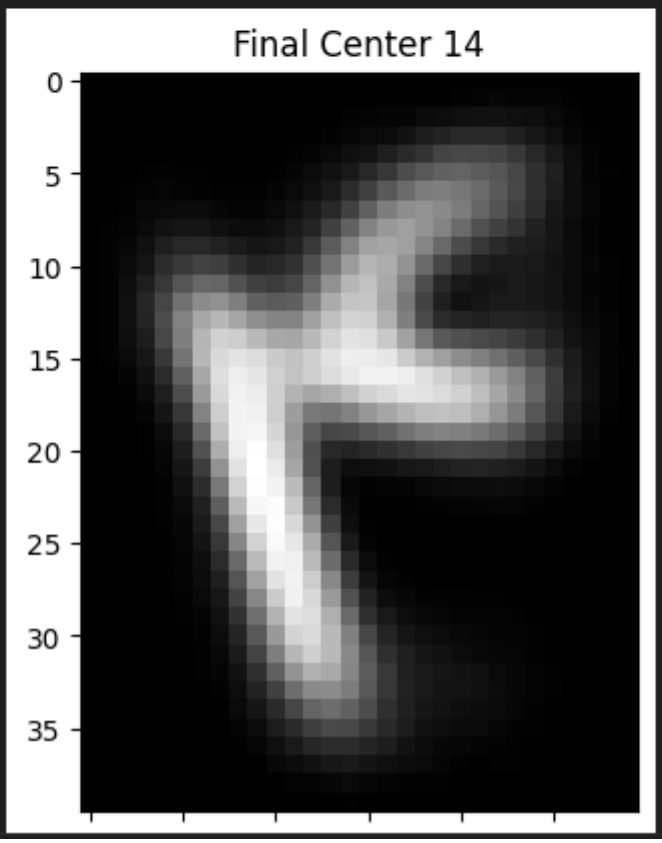

Final Center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25

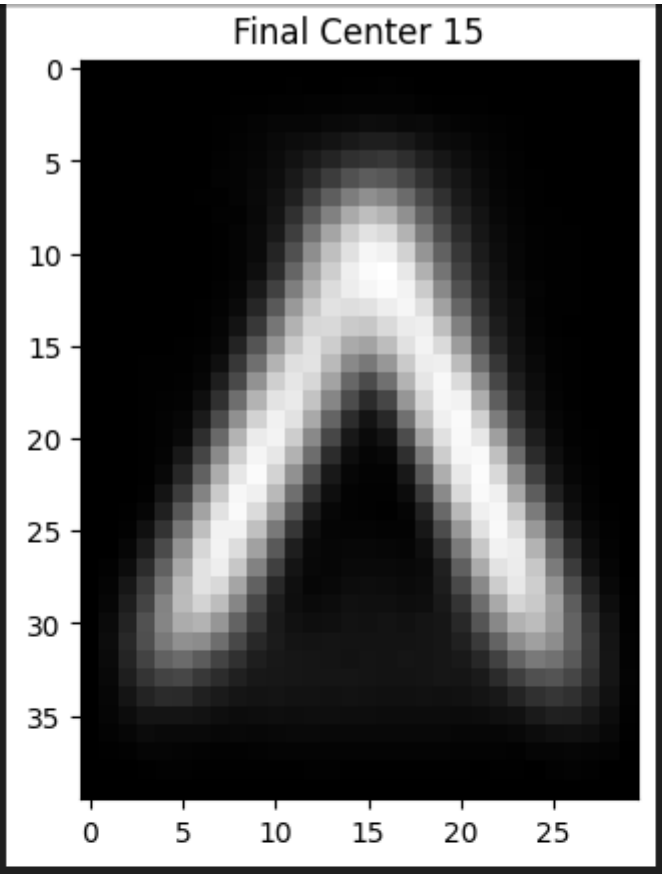

Final Center 10
0
5
10
15
20
25
30
35

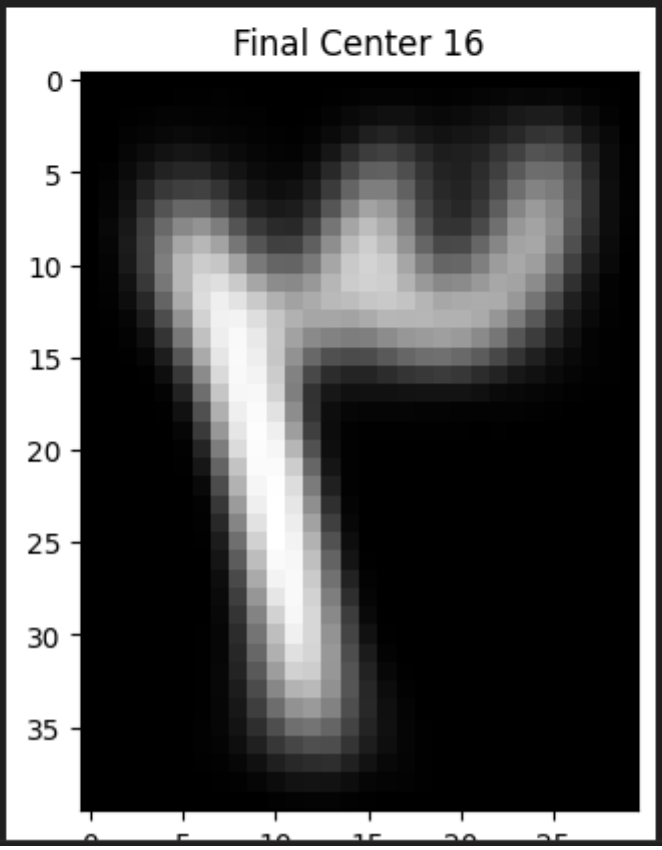

Final Center 11
0
5
10
15
20
25
30
35
0
5
10
15
20
25

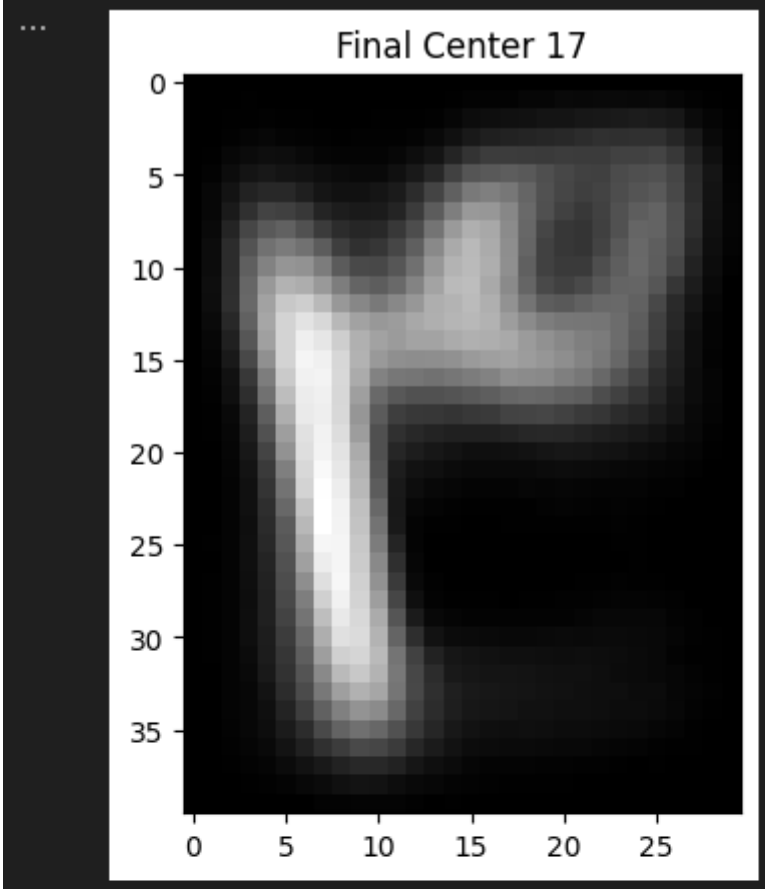

Final Center 12
0
5
10
15
20
25
30
35

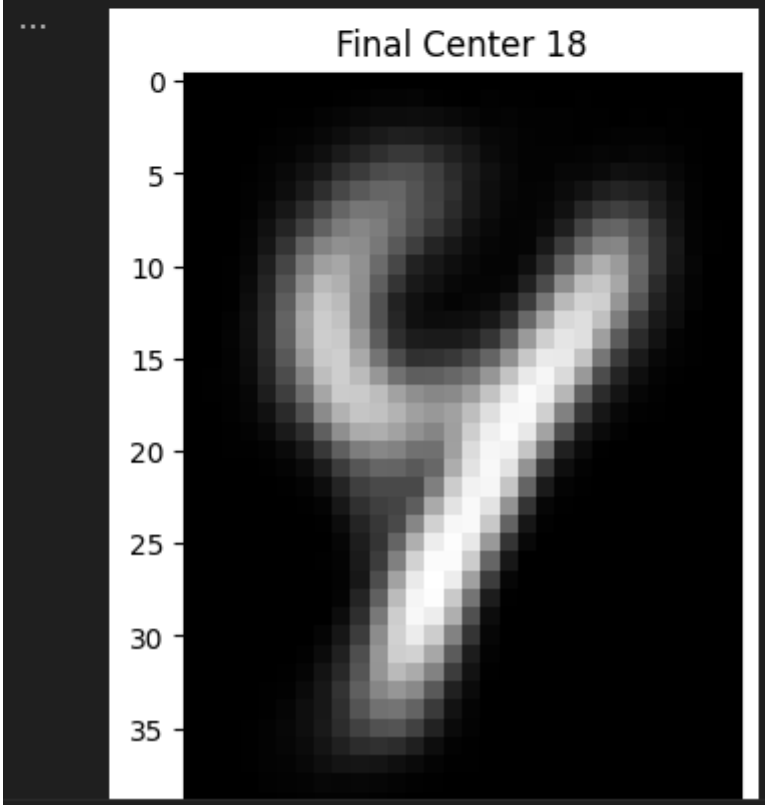

Final Center 13
0
5
10
15
20
25
30
35
0
5
10
15
20
25

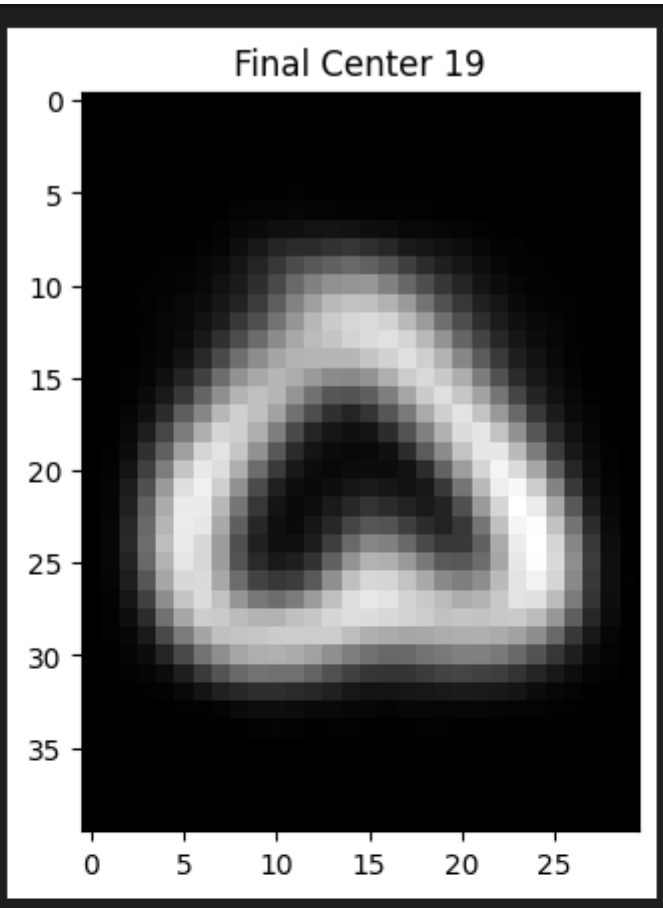

Final Center 14
0
5
10
15
20
25
30
35

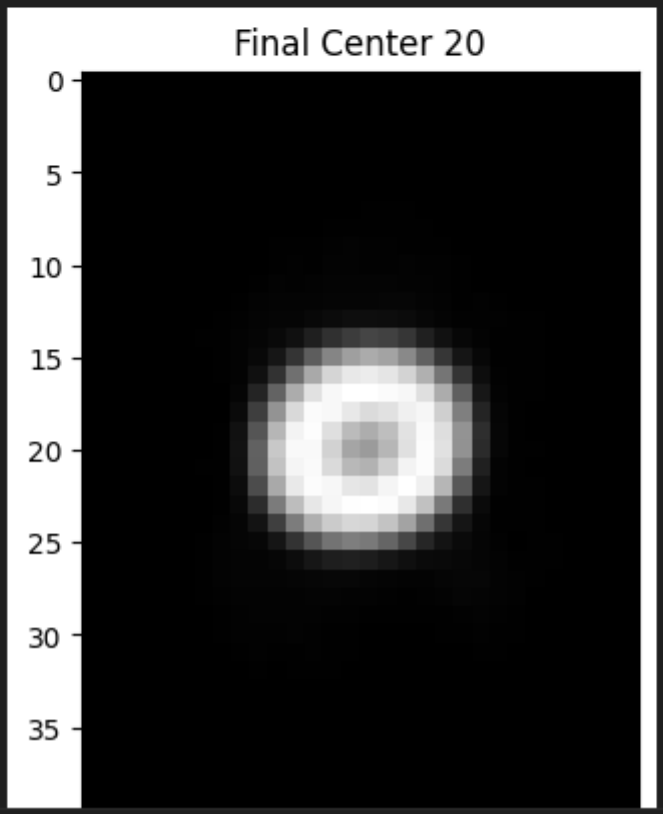

Final Center 15
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 16
0
5
10
15
20
25
30
35

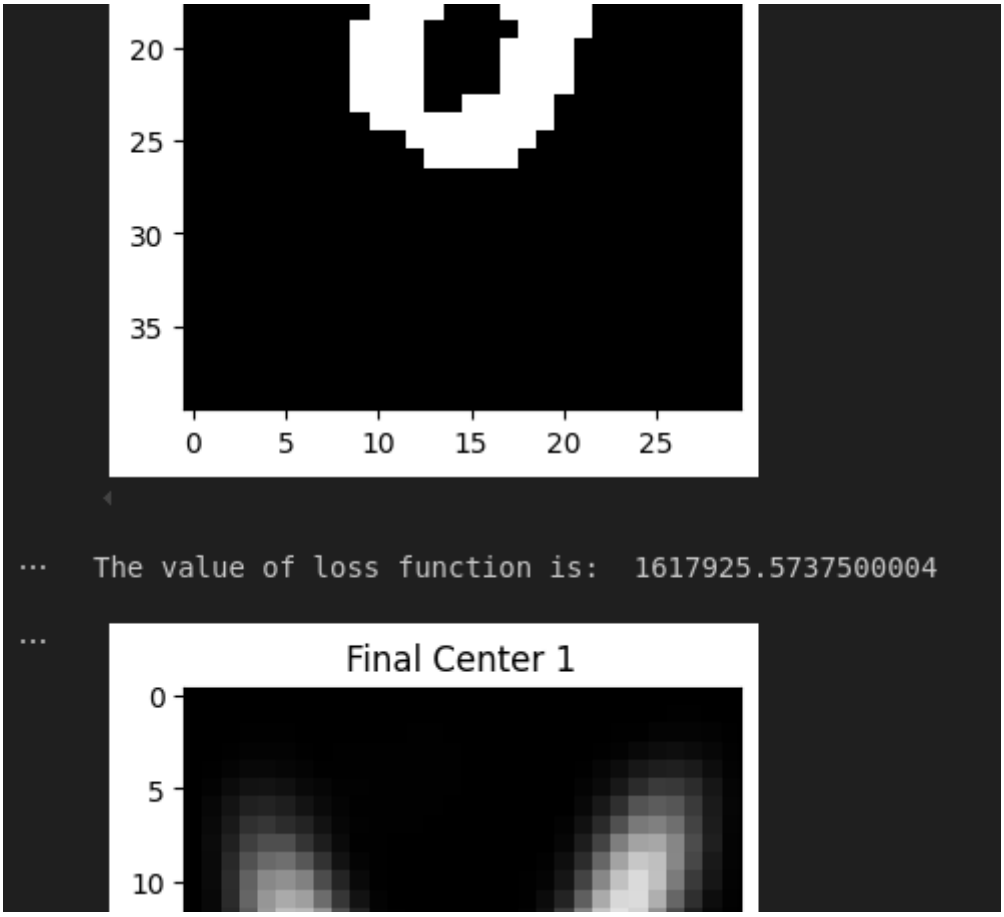

Final Center 17
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 18
0
5
10
15
20
25
30
35


Final Center 19
0
5
10
15
20
25
30
35
0
5
10
15
20
25

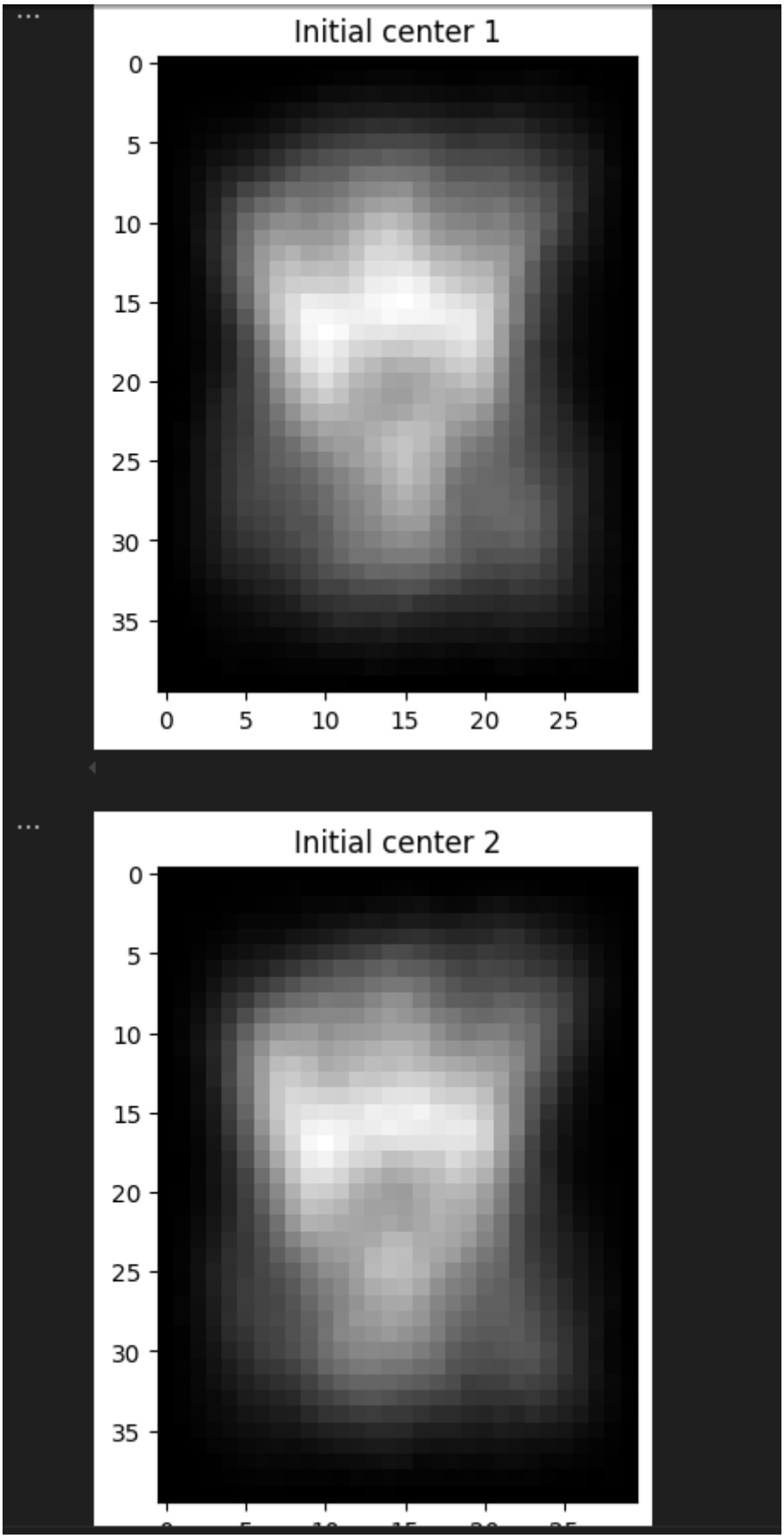

Final Center 20
0
5
10
15
20
25
30
35

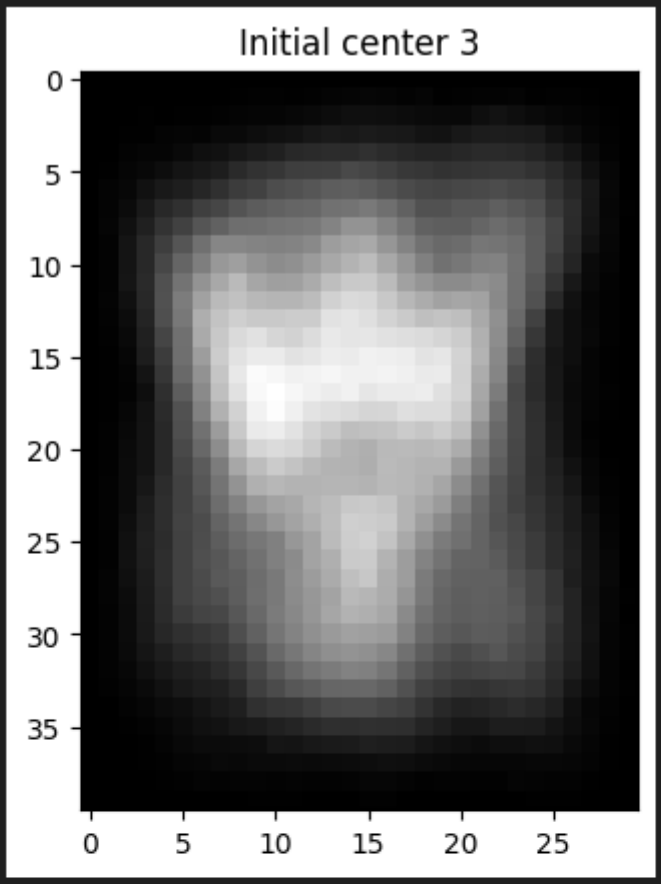
مقدار تابع هزینه

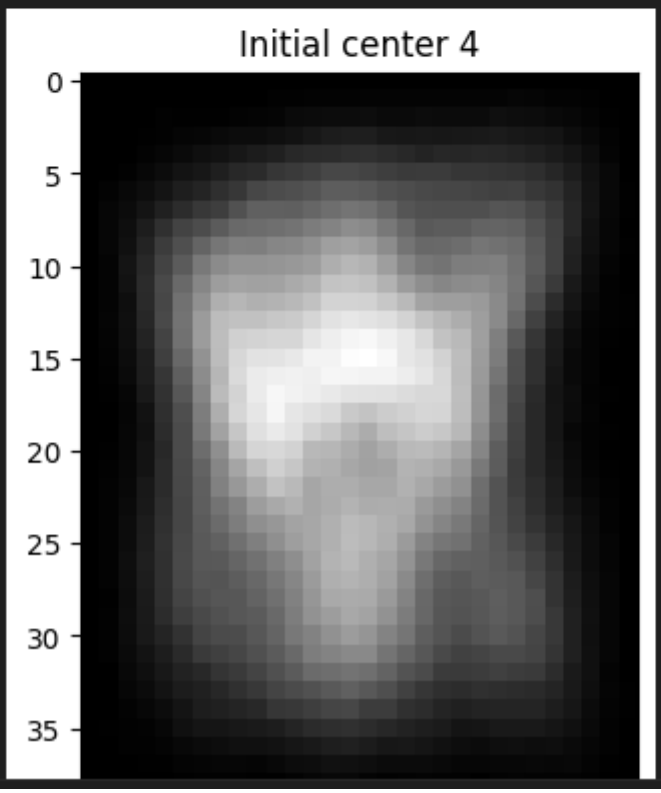


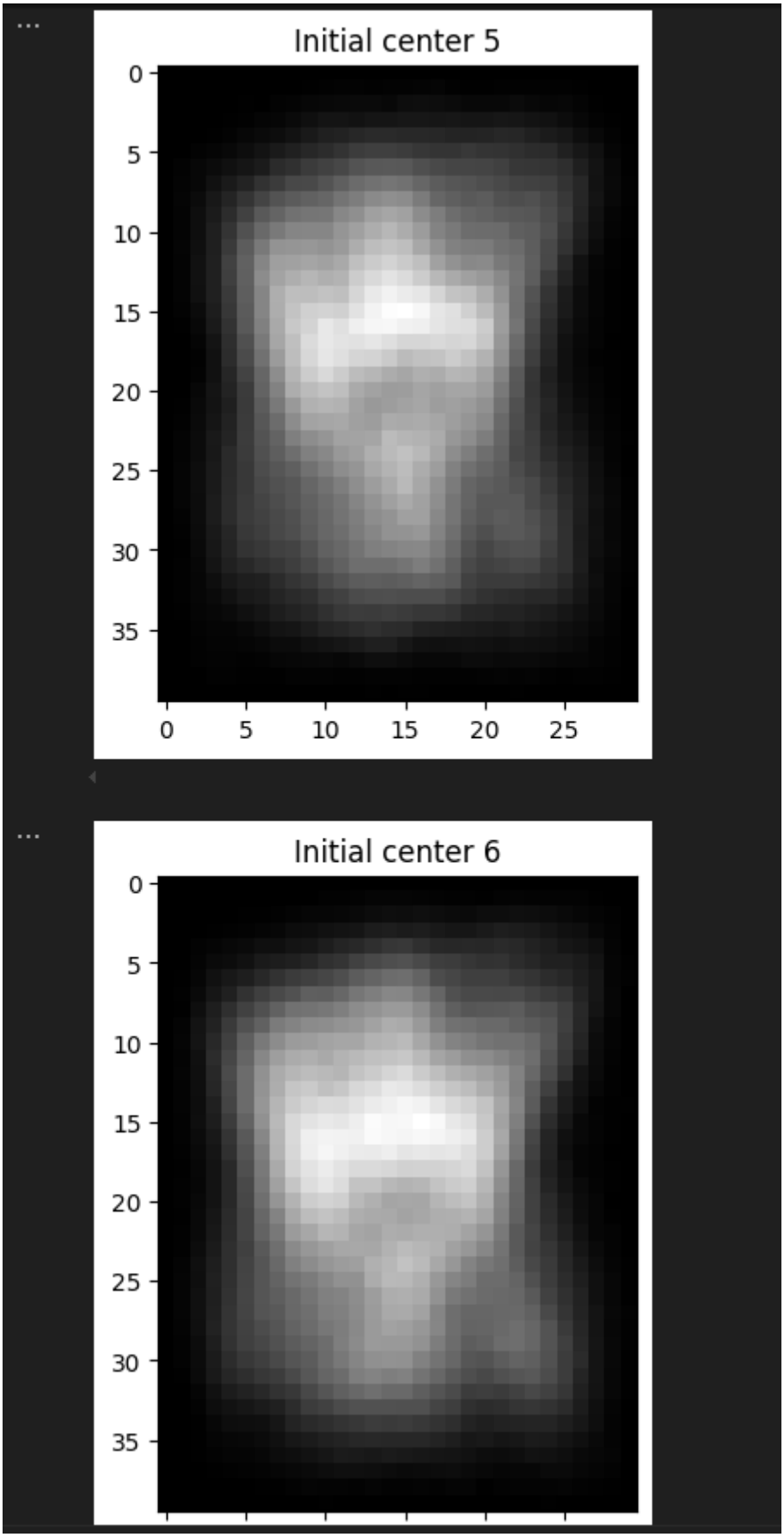
b.III:

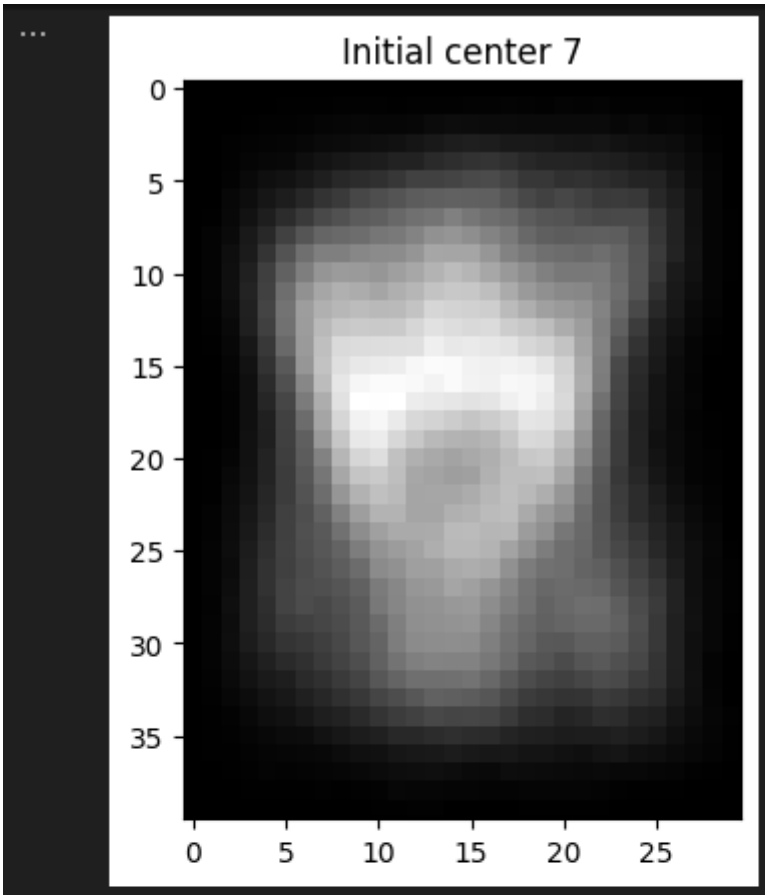
شاخص های اولیه ساخت شده با میانگین هر قسمت

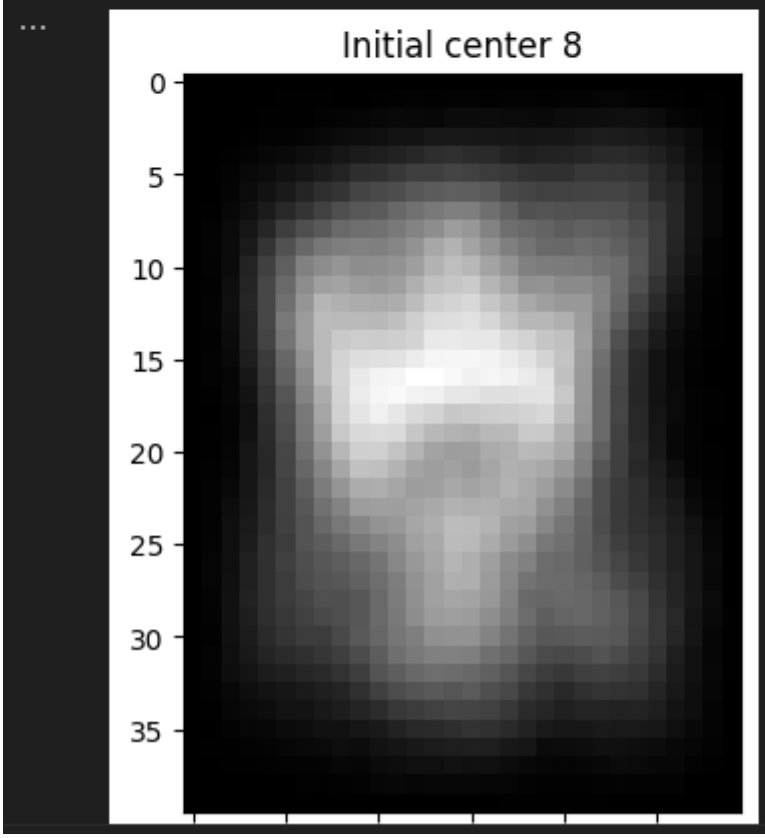


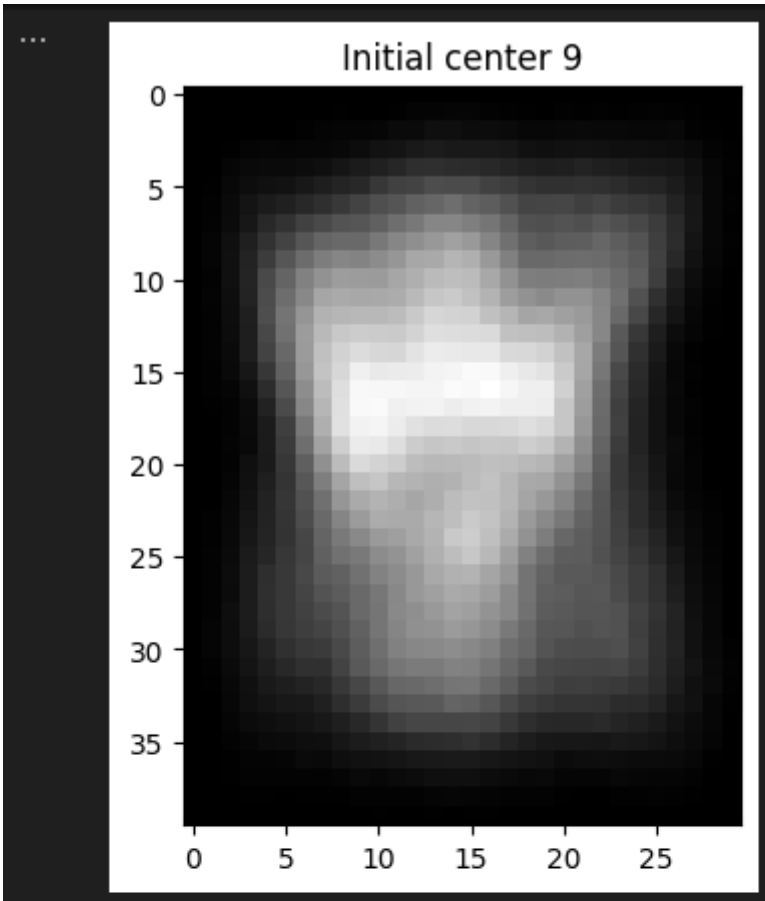

Initial center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25

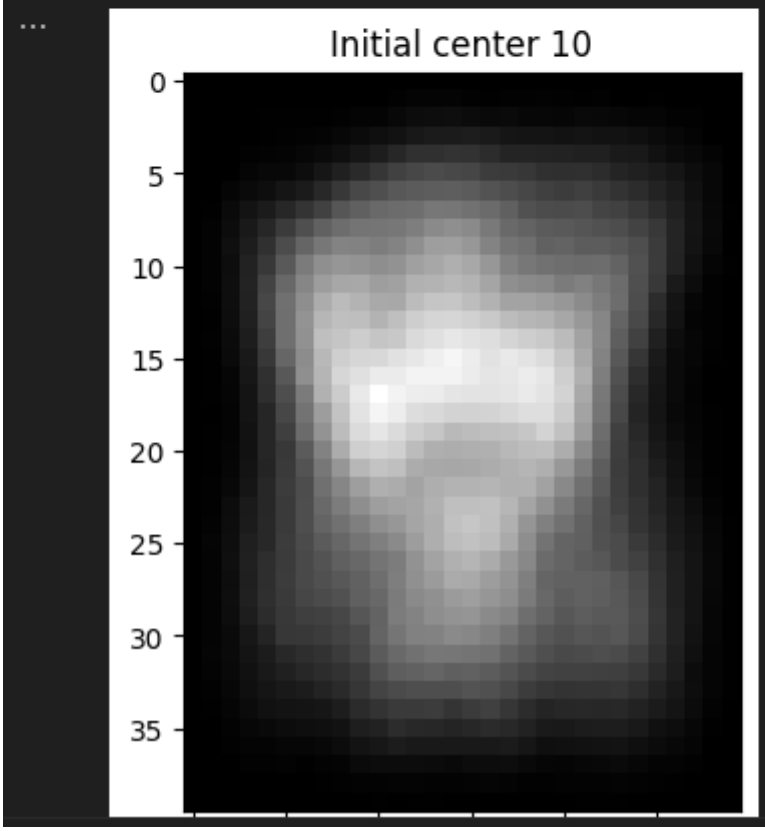

Initial center 4
0
5
10
15
20
25
30
35

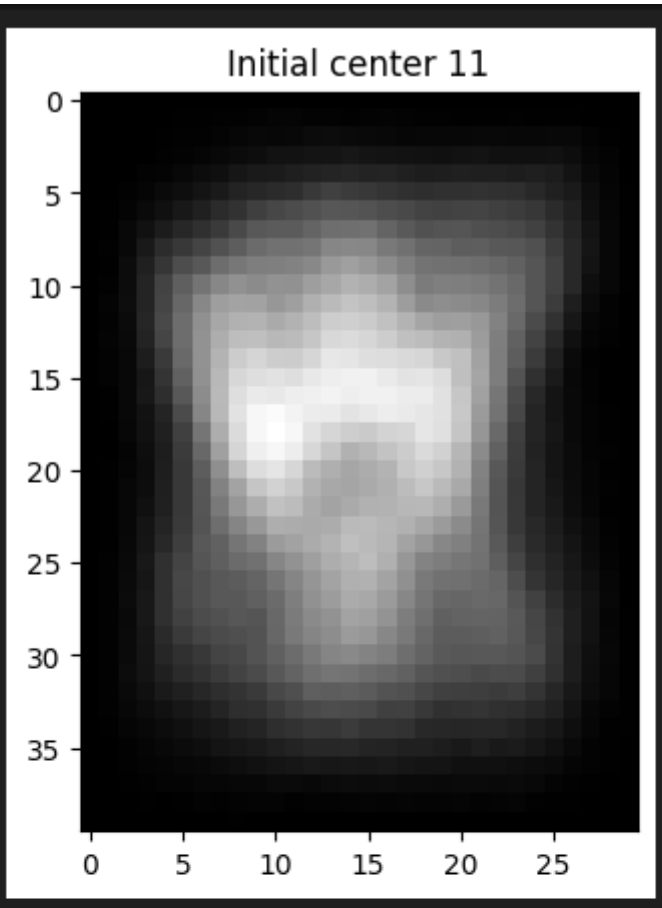

Initial center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Initial center 6
0
5
10
15
20
25
30
35

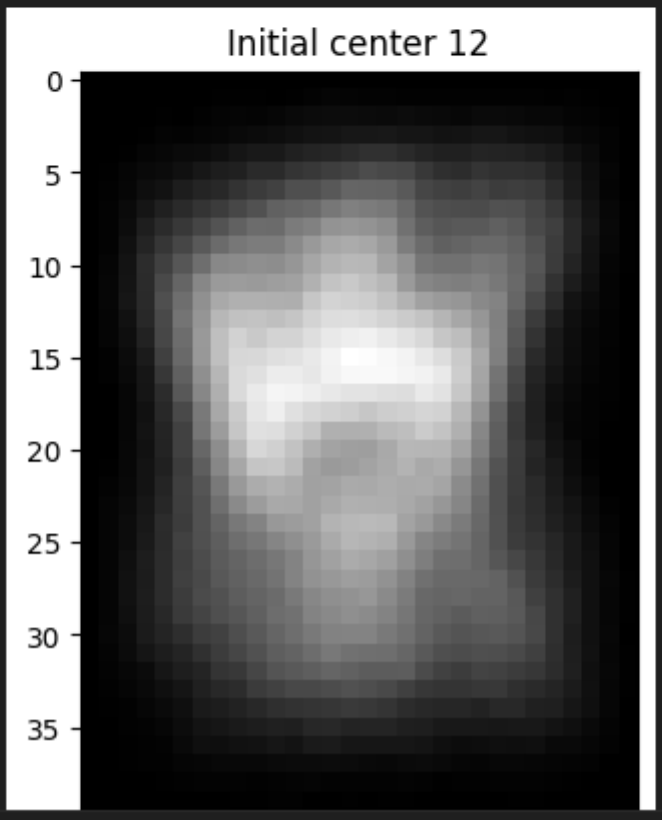

Initial center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25

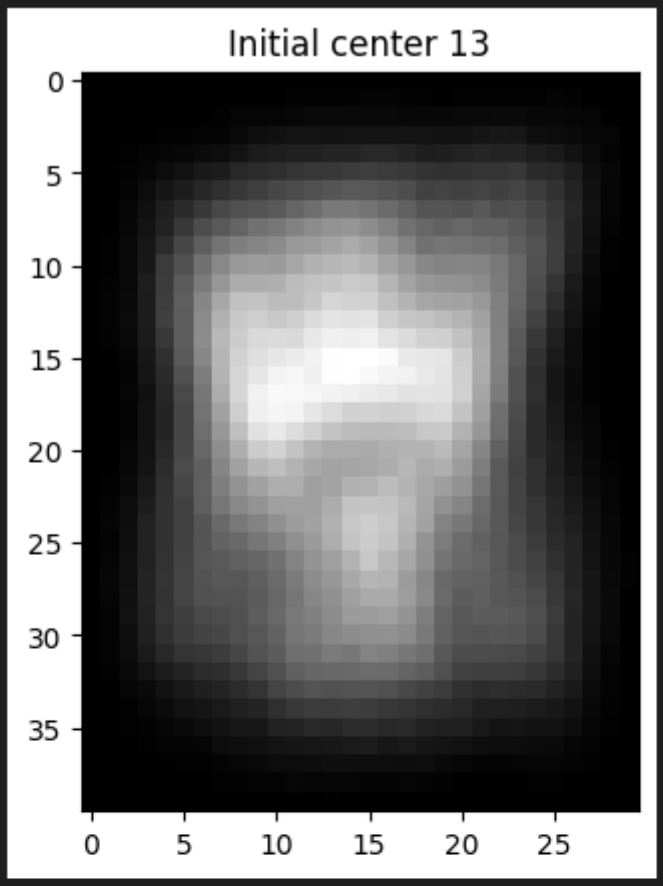

Initial center 8
0
5
10
15
20
25
30
35

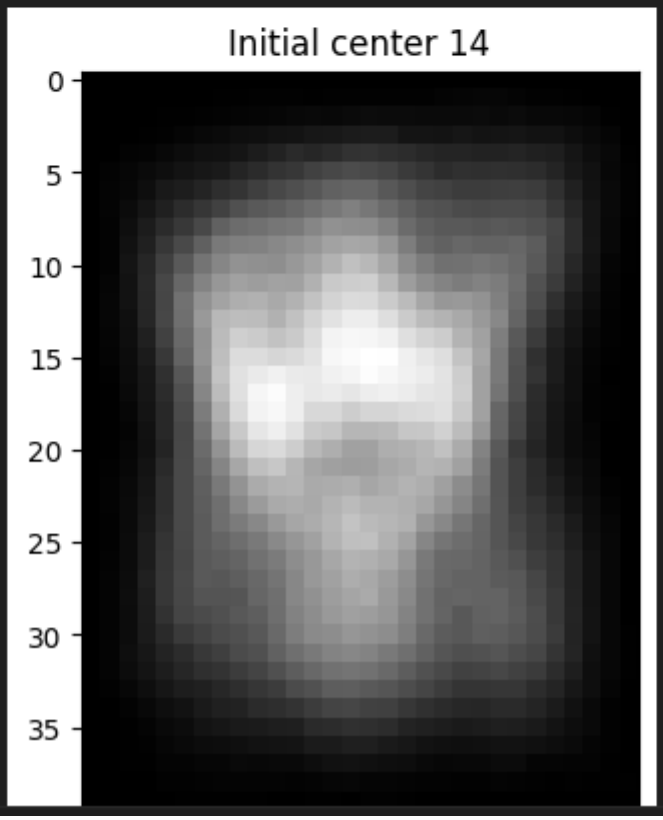

Initial center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25

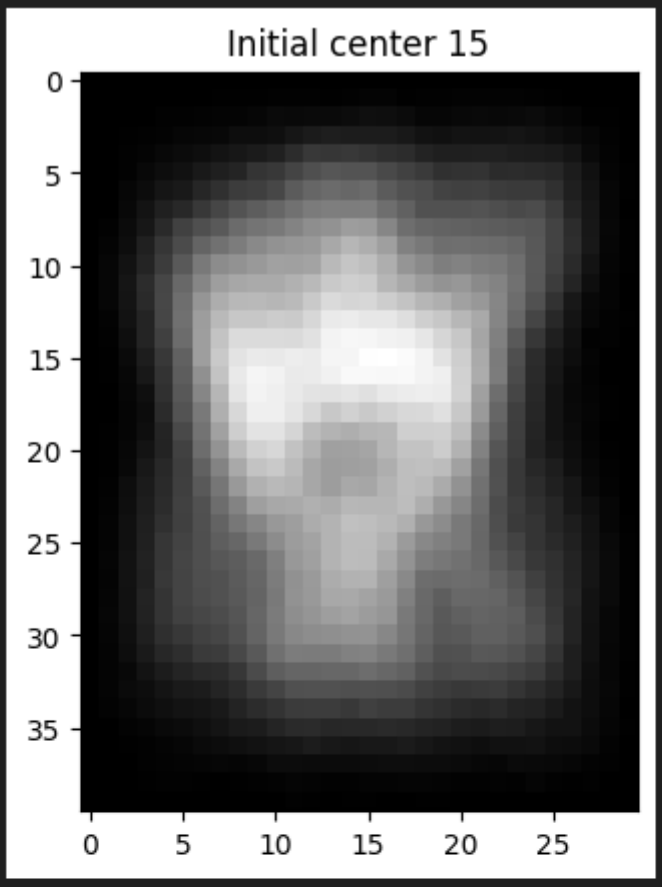

Initial center 10
0
5
10
15
20
25
30
35

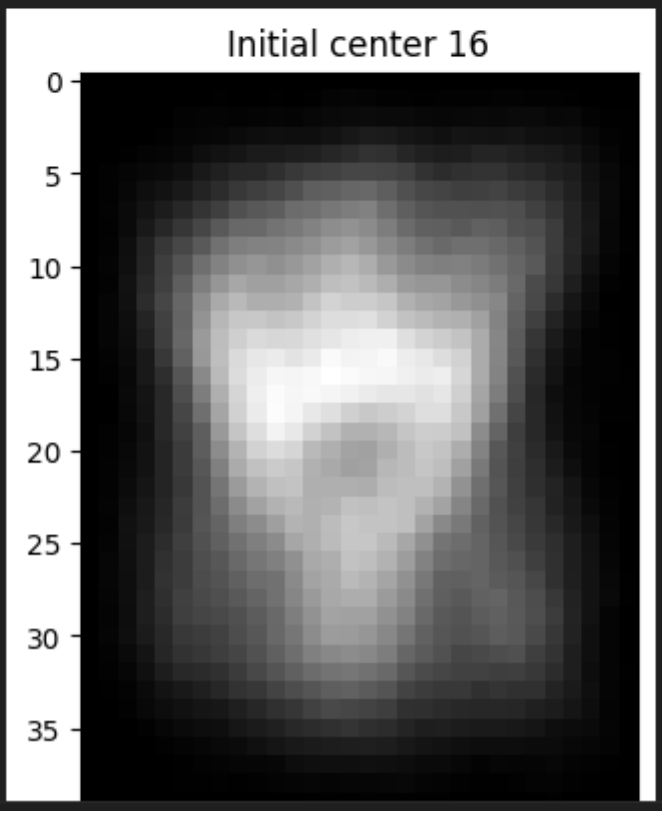

Initial center 11
0
5
10
15
20
25
30
35
0
5
10
15
20
25

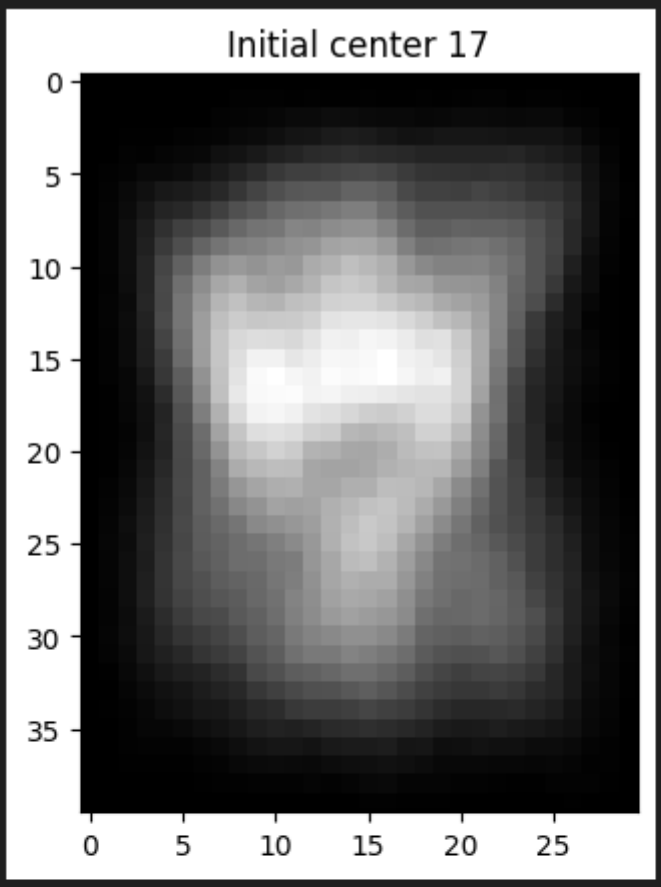

Initial center 12
0
5
10
15
20
25
30
35

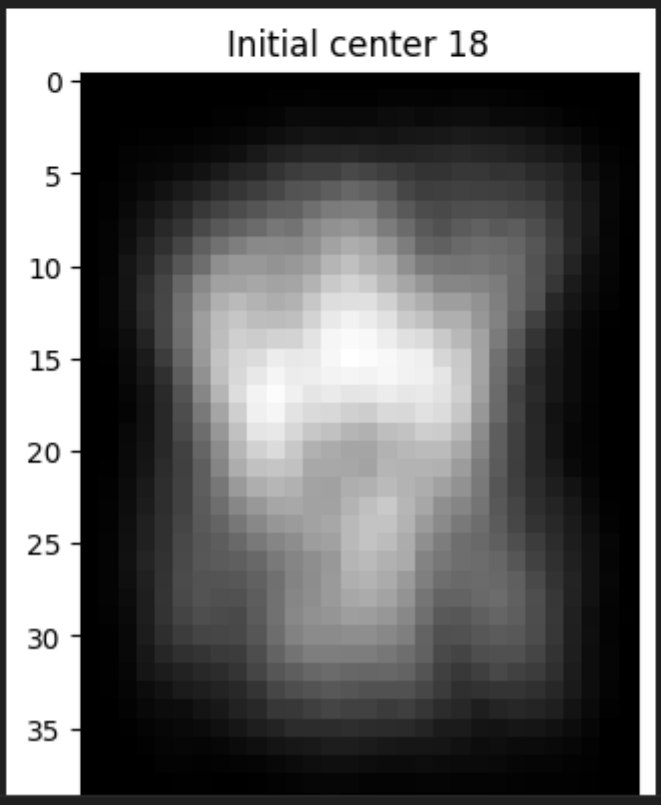

Initial center 13
0
5
10
15
20
25
30
35
0
5
10
15
20
25

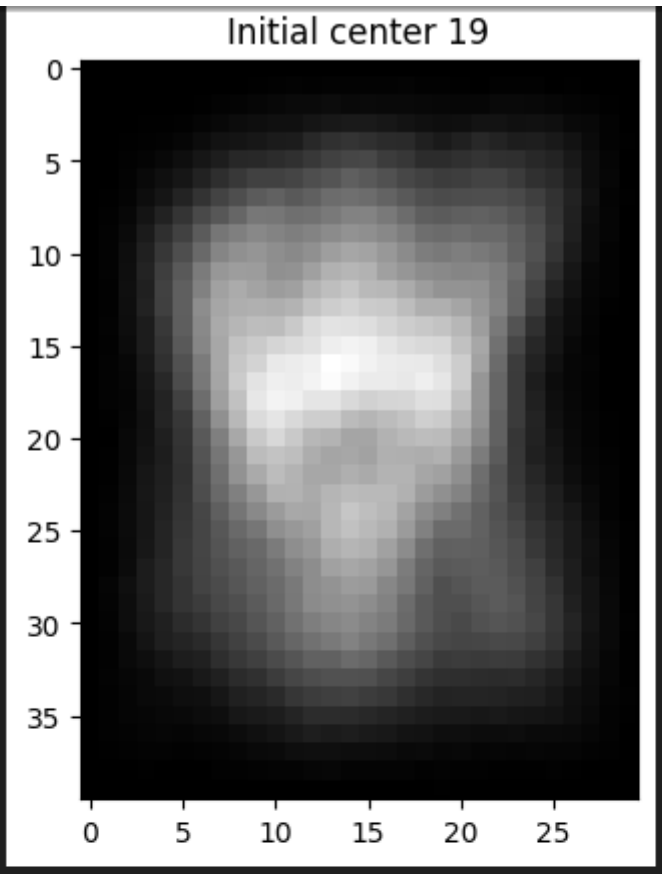

Initial center 14
0
5
10
15
20
25
30
35

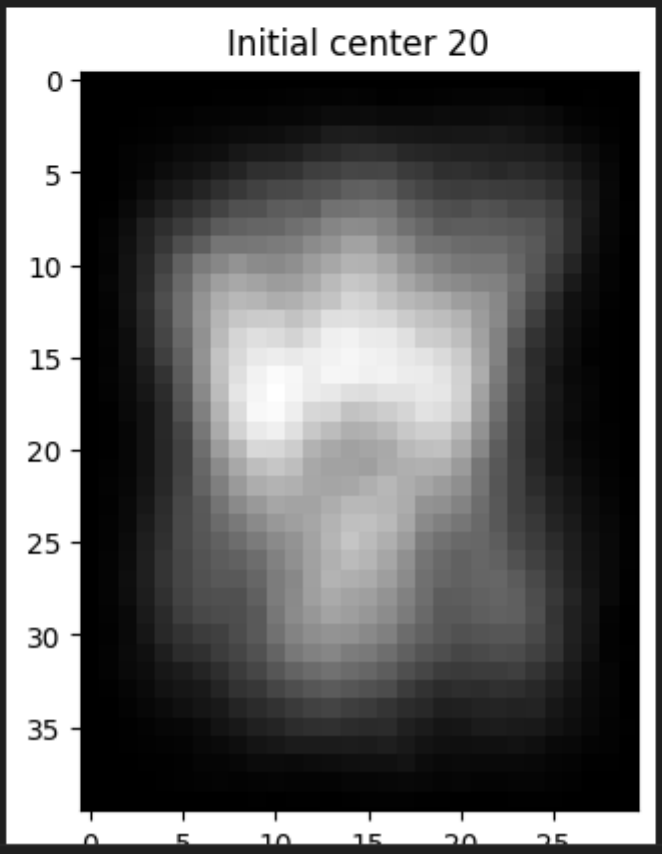

Initial center 15
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Initial center 16
0
5
10
15
20
25
30
35

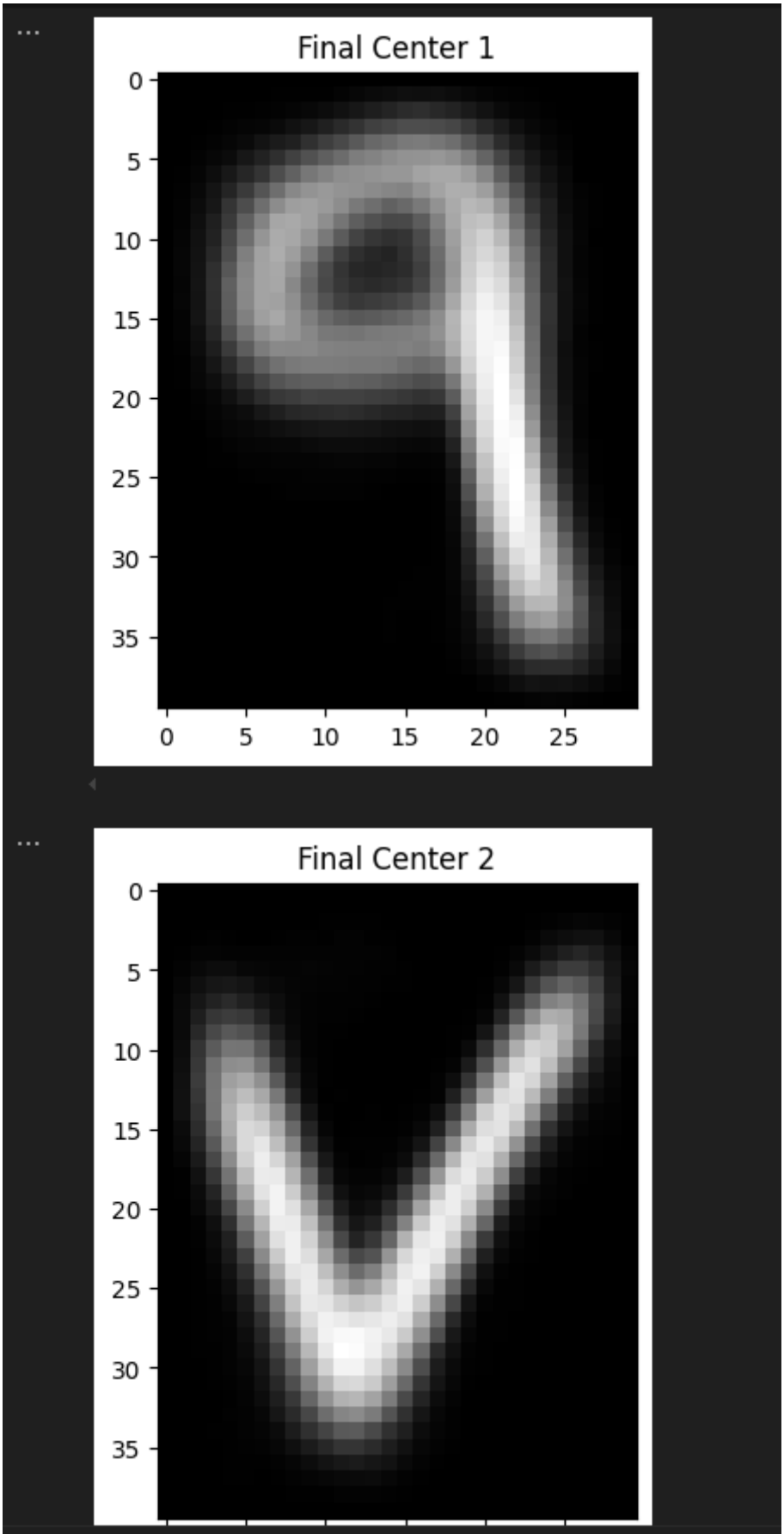

Initial center 17
0
5
10
15
20
25
30
35
0
5
10
15
20
25

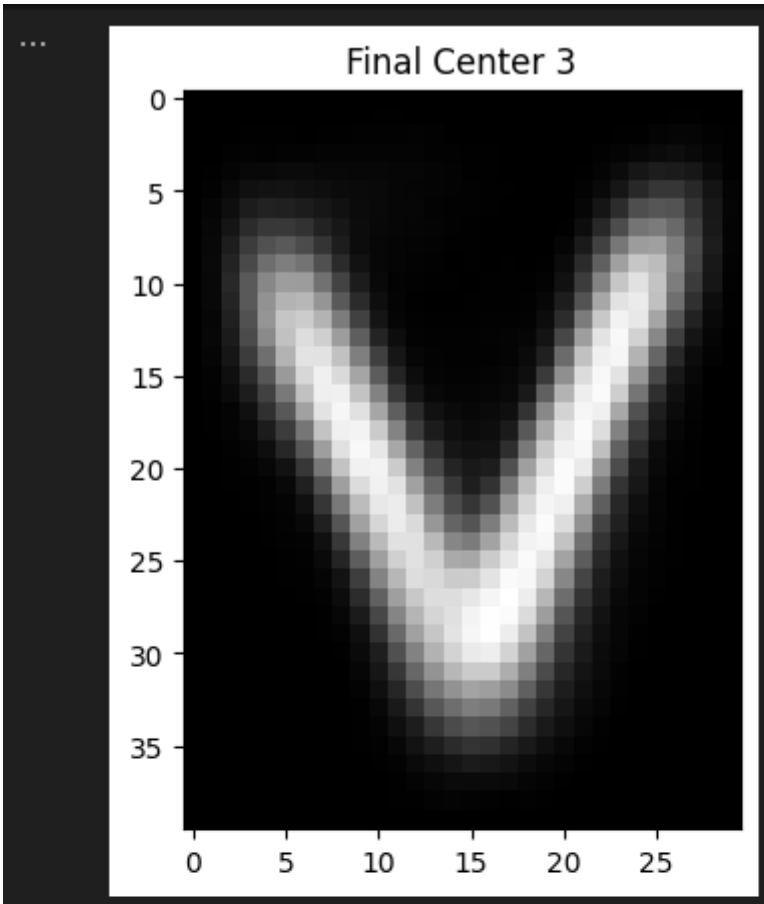

Initial center 18
0
5
10
15
20
25
30
35

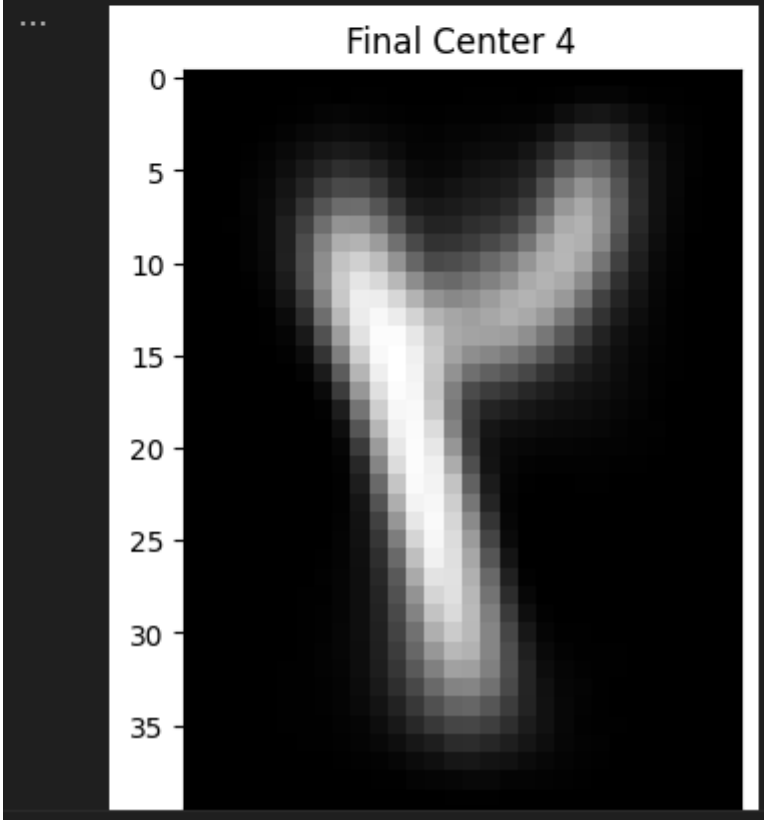

Initial center 19
0
5
10
15
20
25
30
35
0
5
10
15
20
25

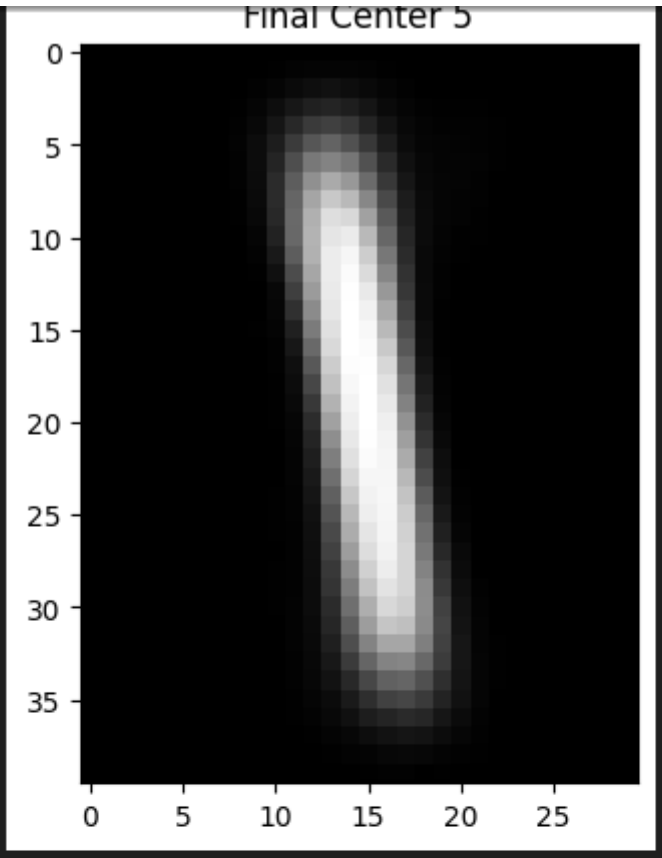

Initial center 20
0
5
10
15
20
25
30
35

شاخص های نهایی

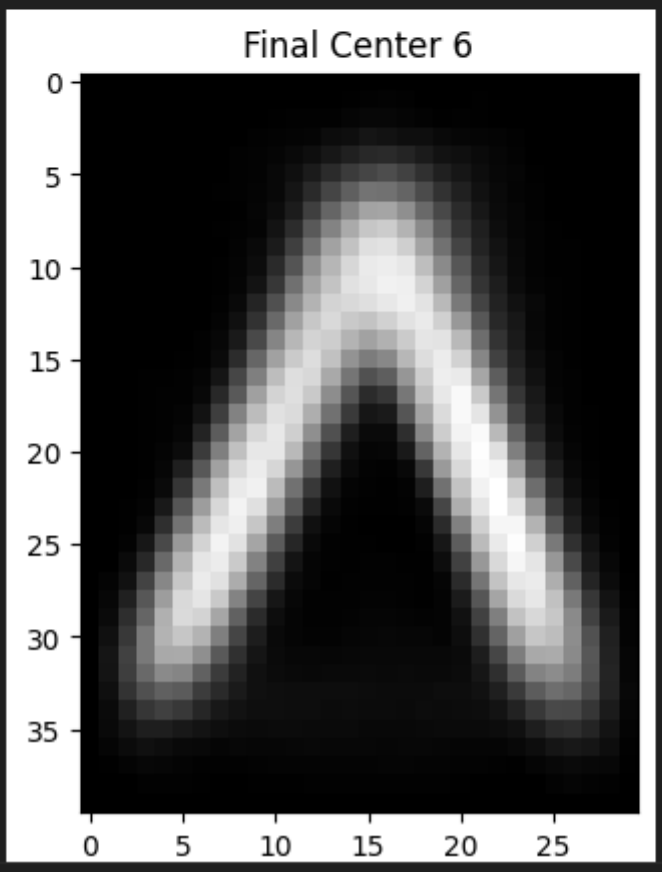


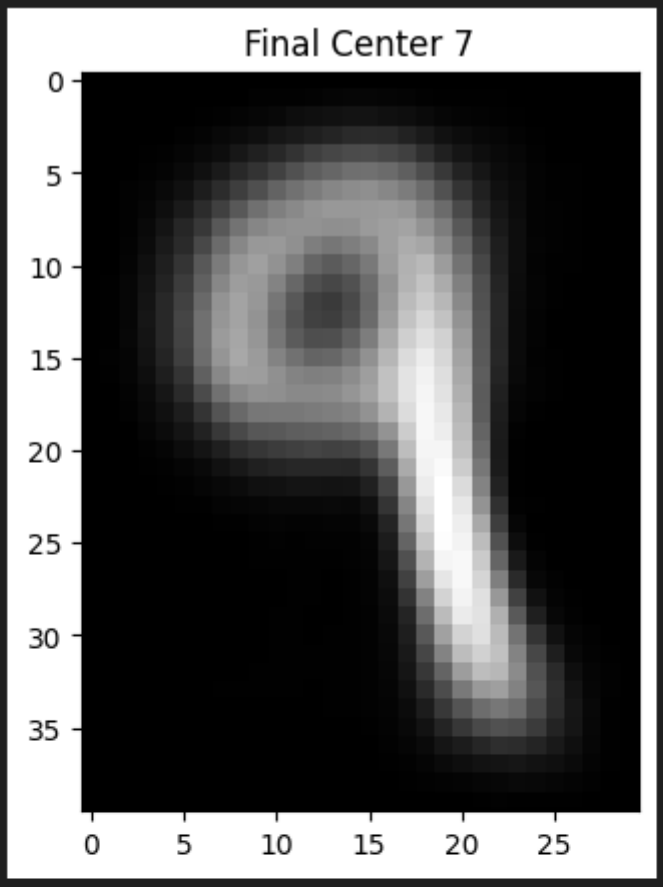

Final Center 3
0
5
10
15
20
25
30
35
0
5
10
15
20
25

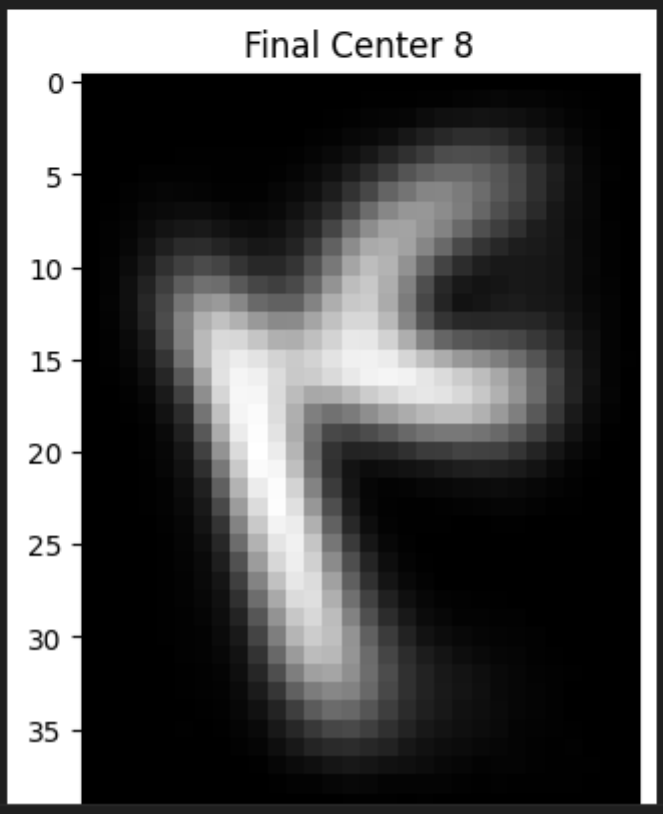

Final Center 4
0
5
10
15
20
25
30
35

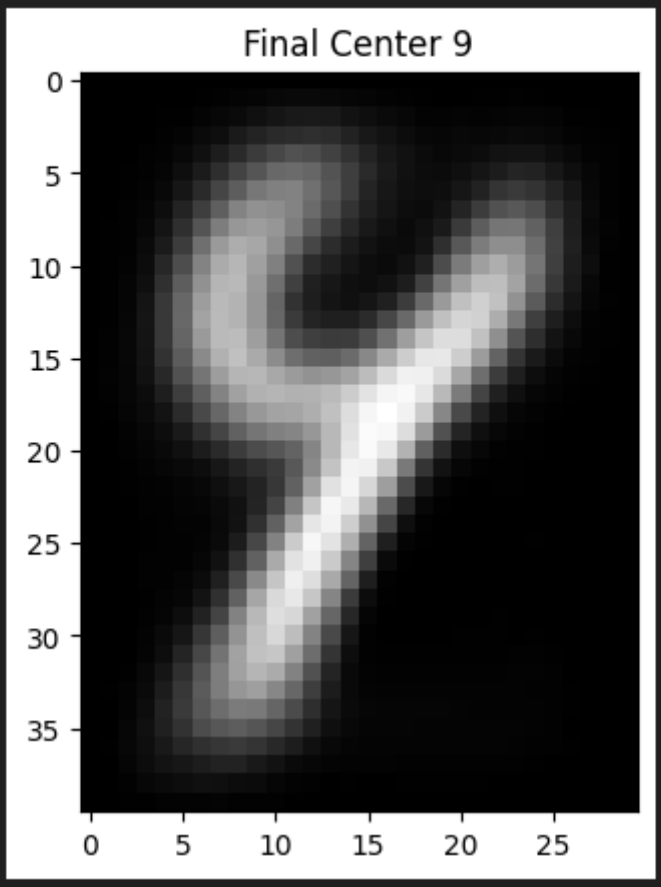

Final Center 5
0
5
10
15
20
25
30
35
0
5
10
15
20
25

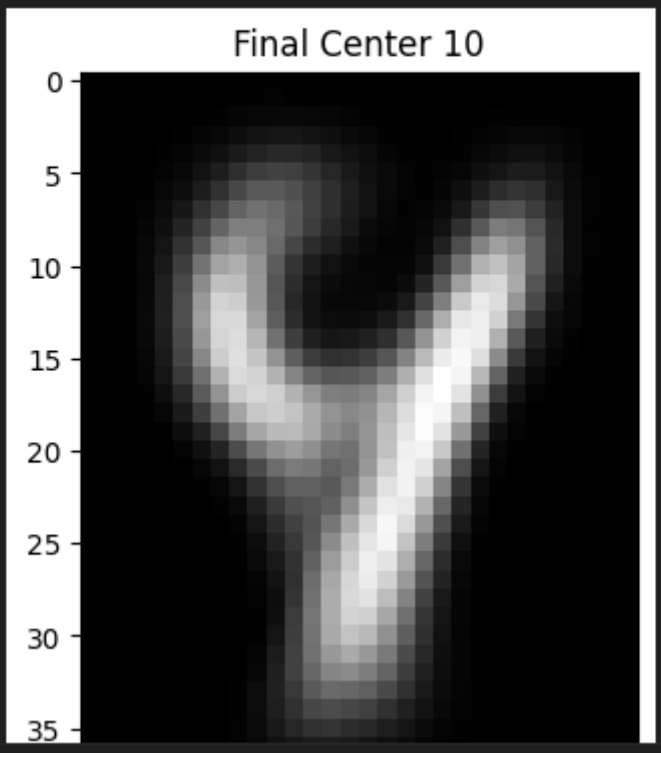
...

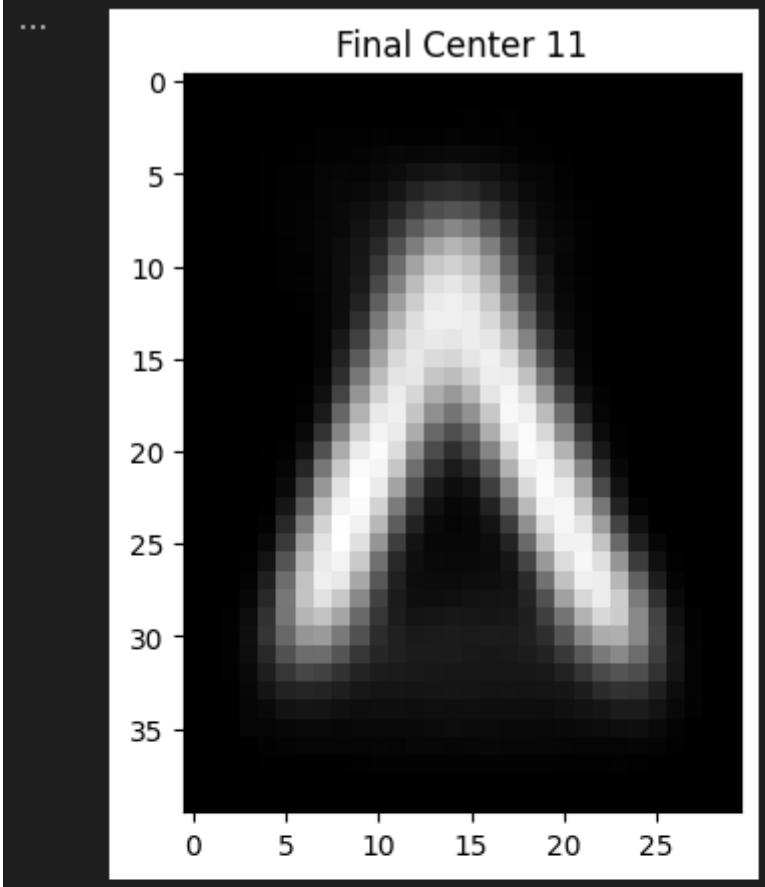

Final Center 6
0
5
10
15
20
25
30
35
0
5
10
15
20
25

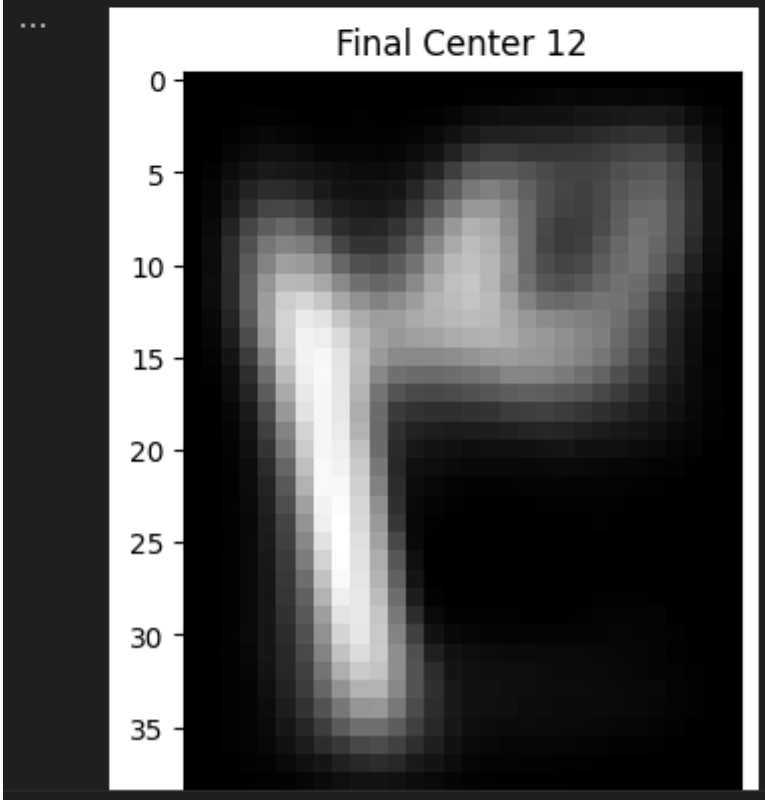

Final Center 7
0
5
10
15
20
25
30
35
0
5
10
15
20
25

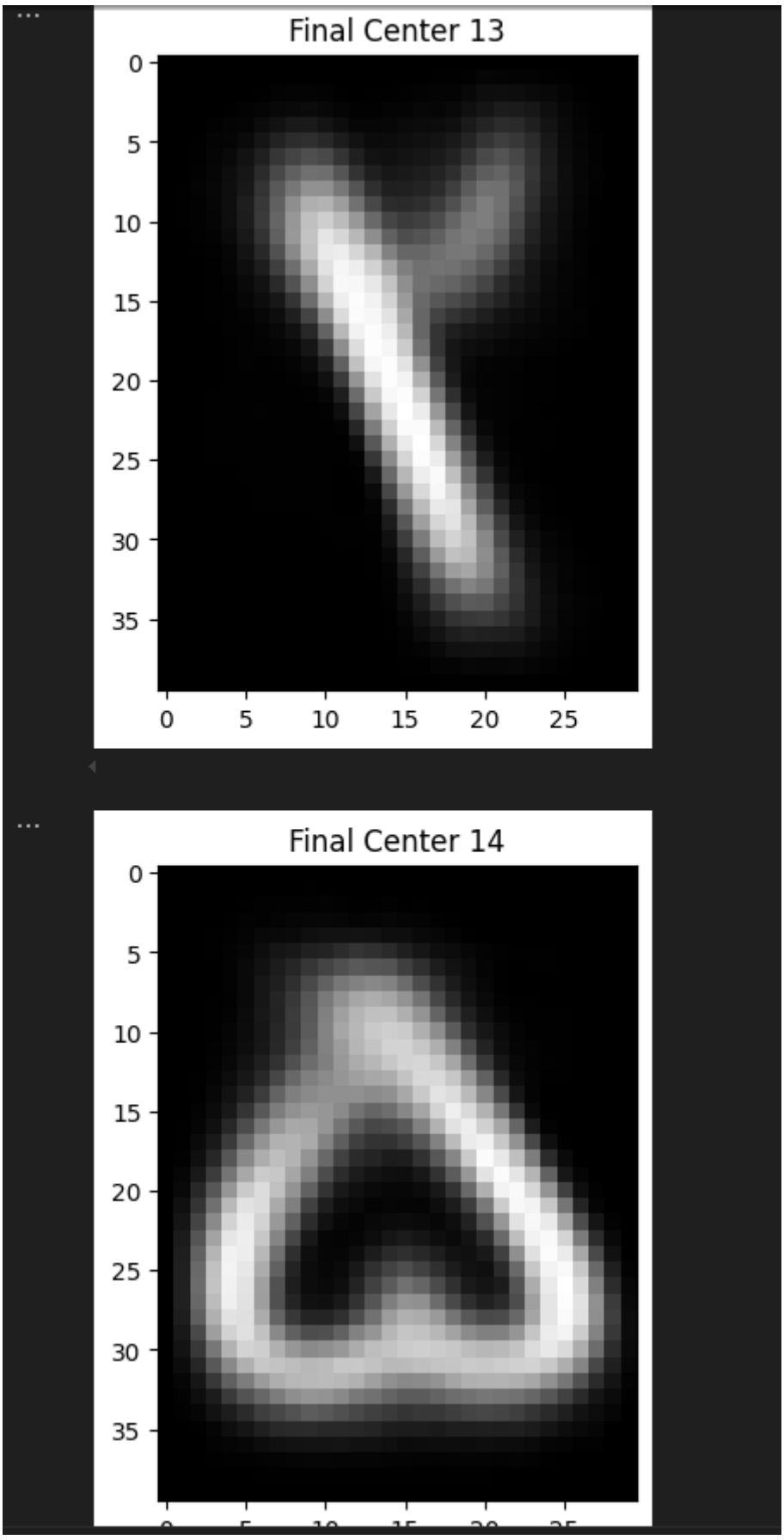

Final Center 8
0
5
10
15
20
25
30
35

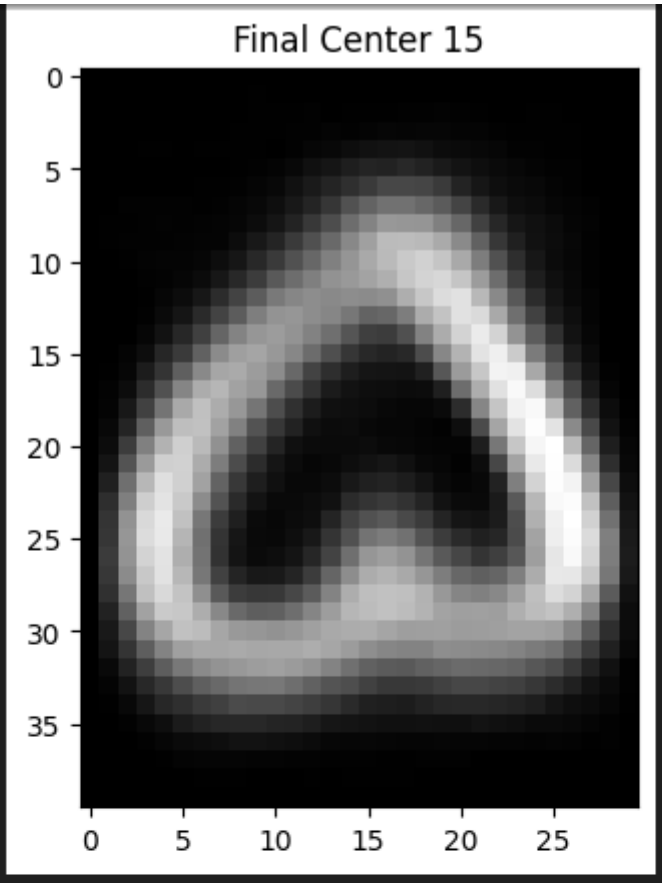

Final Center 9
0
5
10
15
20
25
30
35
0
5
10
15
20
25

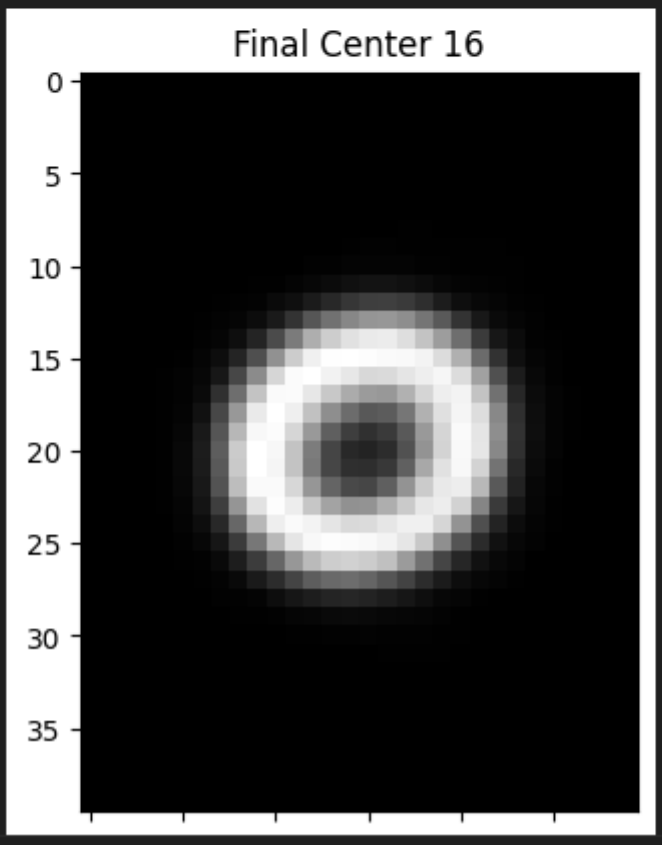

Final Center 10
0
5
10
15
20
25
30
35

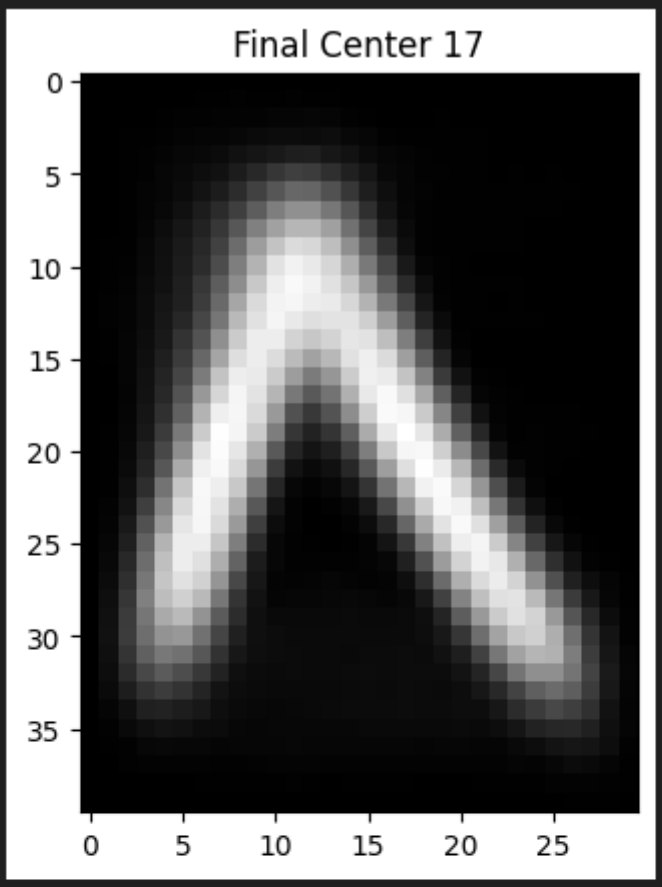

Final Center 11
0
5
10
15
20
25
30
35
0
5
10
15
20
25

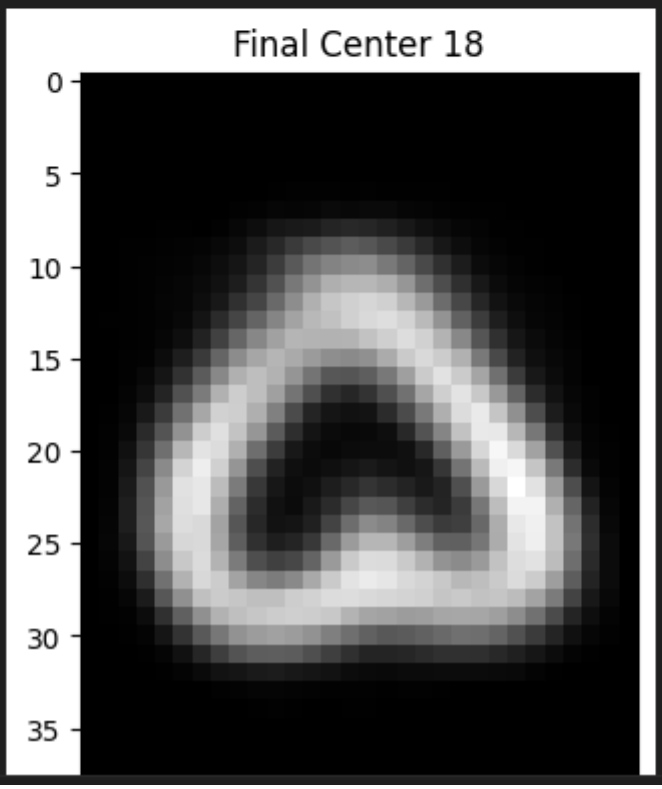

Final Center 12
0
5
10
15
20
25
30
35

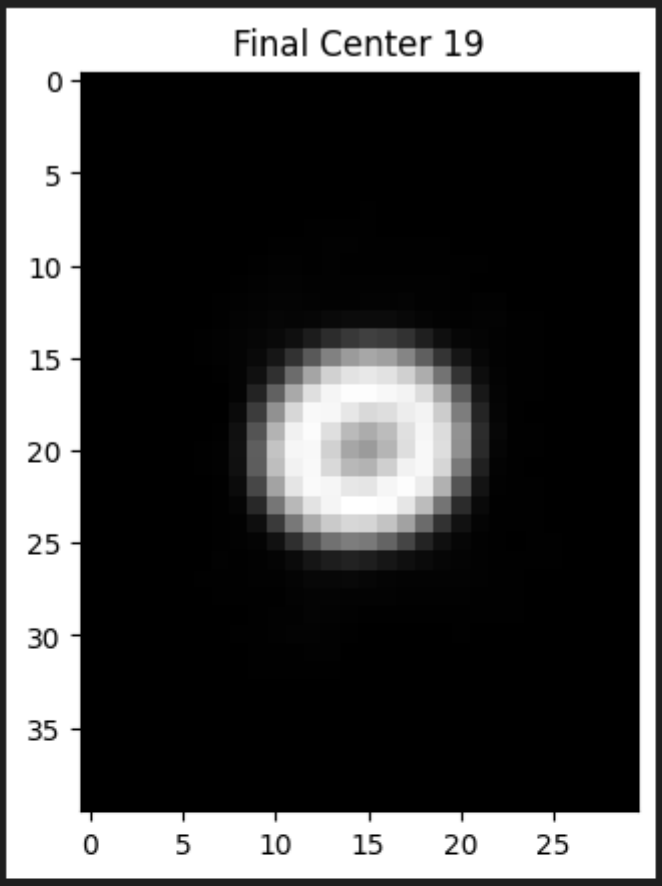

Final Center 13
0
5
10
15
20
25
30
35
0
5
10
15
20
25
Final Center 14
0
5
10
15
20
25
30
35

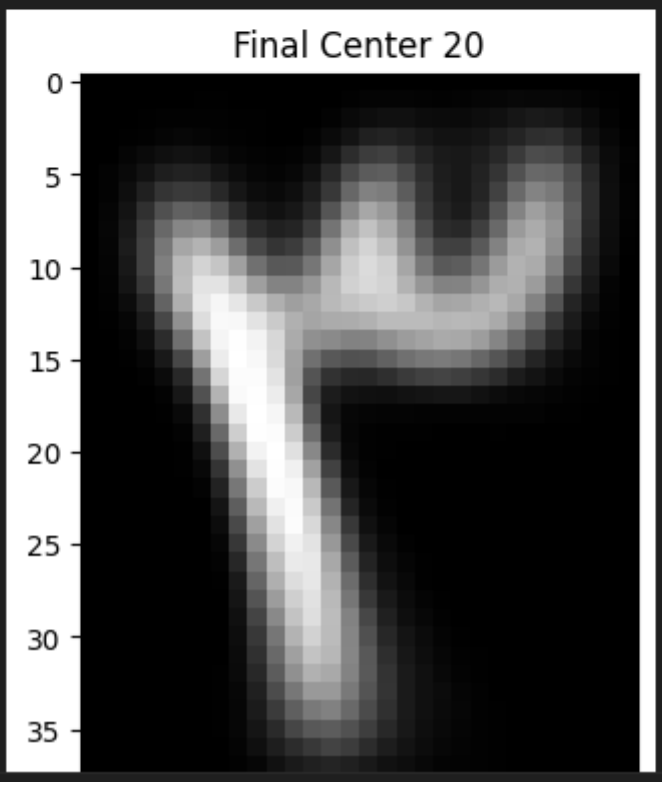

Final Center 15
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 16
0
5
10
15
20
25
30
35

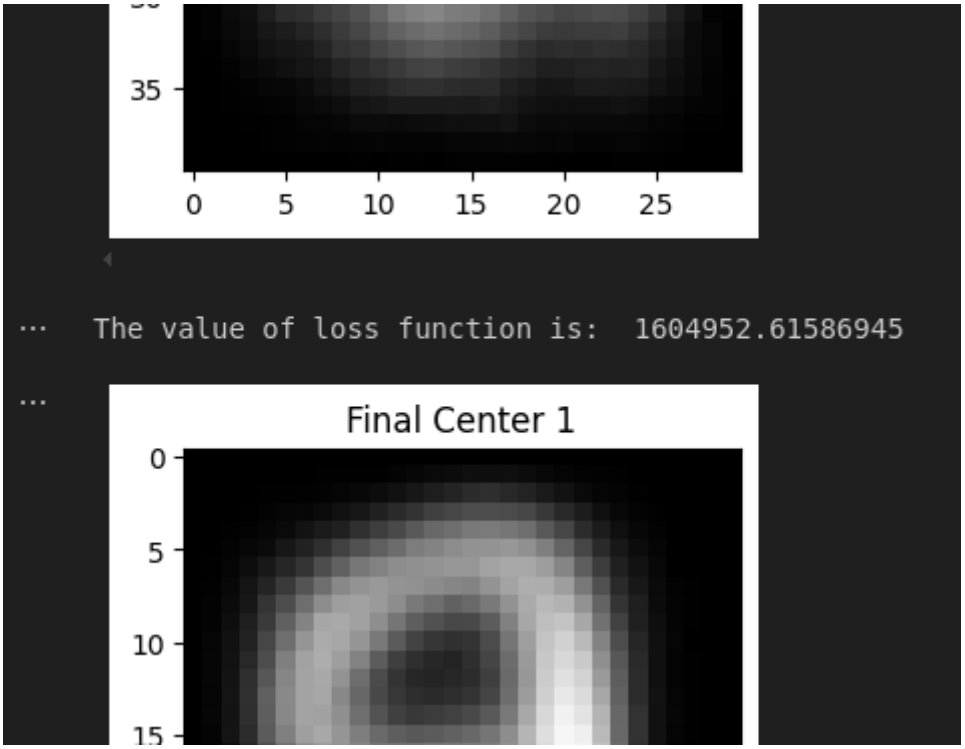

Final Center 17
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 18
0
5
10
15
20
25
30
35


Final Center 19
0
5
10
15
20
25
30
35
0
5
10
15
20
25


Final Center 20
0
5
10
15
20
25
30
35

مقدار تابع هزینه



تحلیل

از این دو قسمت نتیجه میگیریم که انتخاب مناسب شاخص اولیه میتواند منجر به به دست آوردن شاخص های نهایی مناسب تر و دقیق تر و خوشه بندی بهتر شود ولی تفاوت زیادی در میزان تابع هزینه ایجاد نمی کند
اگر تعداد خوشه ها را زیاد کنیم میزان تابع هزینه کم می شود که طبیعی است در این مورد برای بعضی از اعداد چند خوشه در نظر گرفته شود

در کذ برای خواندن داده ها از توابعی کتابخانه پانداس و برای کار با آرایه ها از کتابخانه نامپای هم چنین از الگوریتم خوشه بندی کتابخانه سایکیت لرن استفاده شده است

c:

Data value	a.I	a.II	a.III	a.IV	b.I	b.II	b.III
0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	2	2
4	4	3	4	4	4	3	4
4	4	8	8	4	4	8	3
5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5
6	2	2	3	6	6	6	6
6	9	9	9	9	9	9	9
7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9

در کل به نظر میرسد استفاده از شاخص های اولیه که به صورت رندوم ساخته شدند دقت خوبی دارد و دقت آن در بعضی موارد نسبت به استفاده از میانگین و انتخاب رندوم از بین داده ها بهتر است هم چنین در کل در مواردی که تعداد خوشه ها ۲۰ تاست دقت کمی بیشتر است

دستورات و توابع استفاده شده

pandas: read_csv() for reading data from .txt file

matplotlib.pyplot: imshow() , title(), show() for displaying an image

numpy: np.random.randint() , np.random.choice for generating or selecting random items
reshape() , np.array_split() for working with arrays

sklearn.cluster: Kmeans , Kmeans.fit() , Kmeans.predict() ,
Kmeans.inertia_ ,Kmeans.cluster_centers_ for clustering