



دانشگاه صنعتی امیرکبیر
دانشکده مهندسی کامپیوتر

پروژه درس شبکه‌های عصبی گزارش عملکرد شبکه کوهون (SOM)

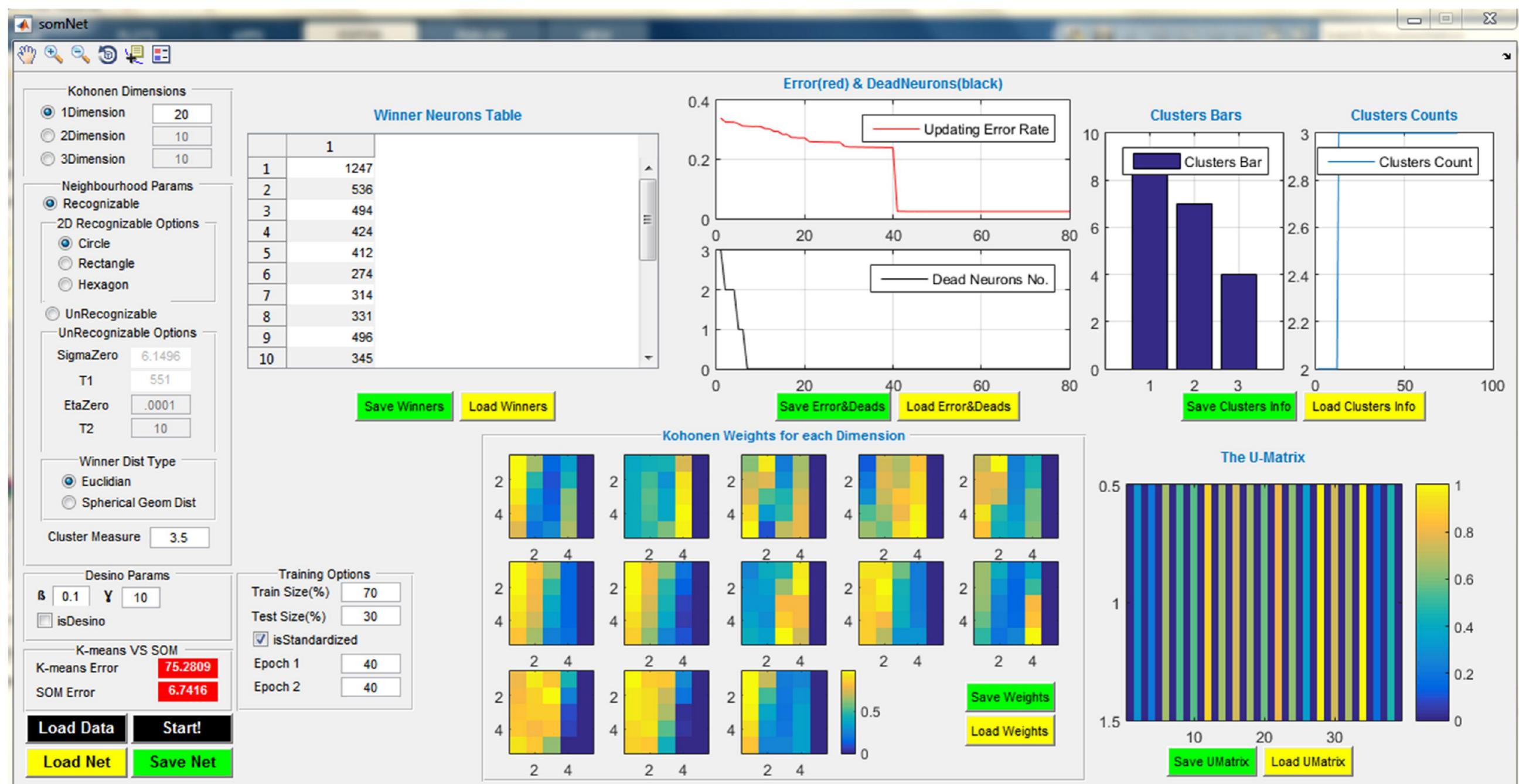
دانشجو:
سید احمد نقوی نوزاد

استاد:
دکتر صفابخش

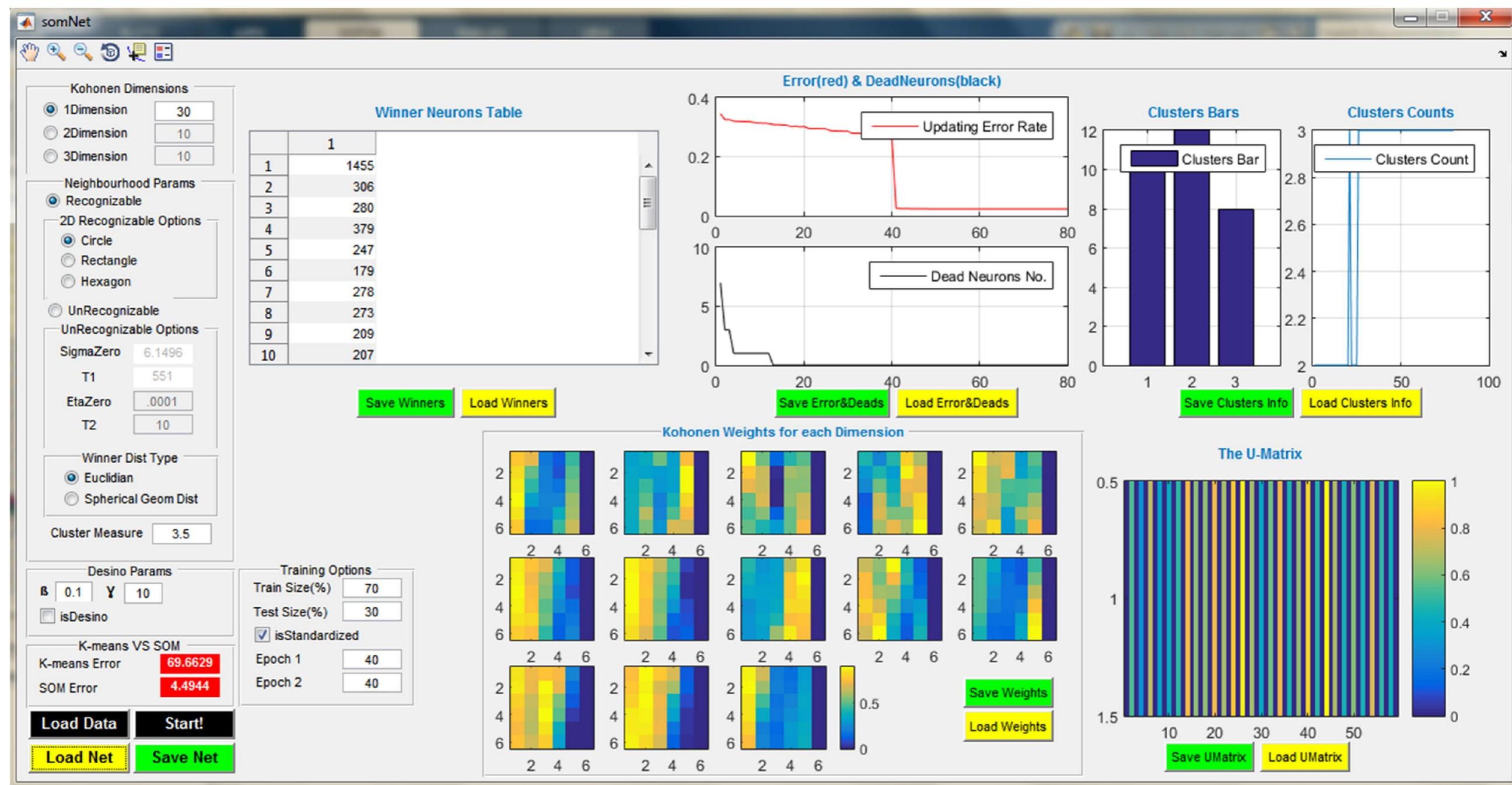
۱. بررسی معماری‌های مختلف در حالت‌های یکبعدی، دوبعدی و سهبعدی

آزمایش اول (حالت یکبعدی)

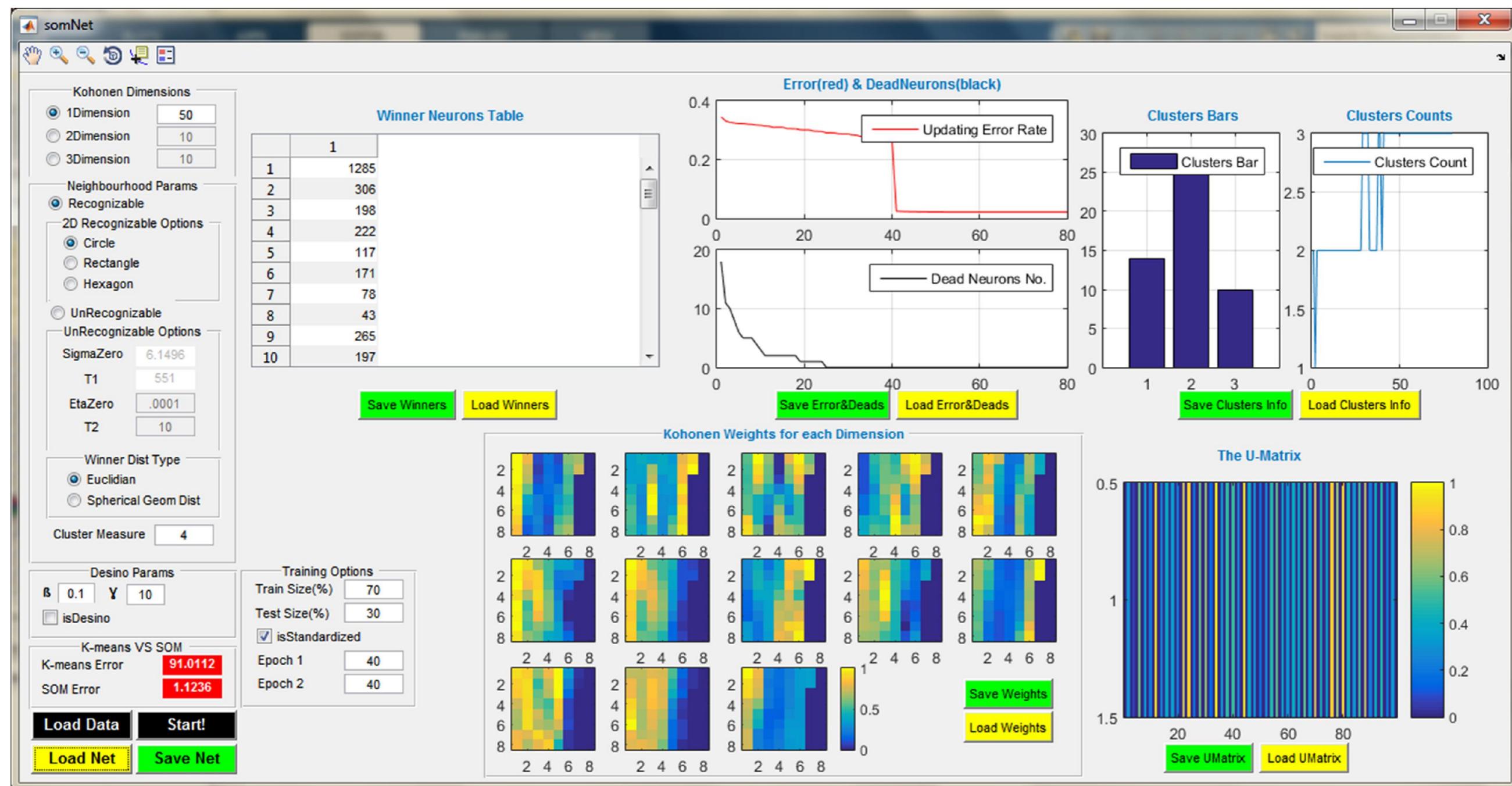
Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 01	[20]	1-D -> Radial	SigmaZero T1 EtaZero T2	Euclidian	3.5	No β γ	.02	0	3	70	30	40	40	Yes	k-means: 75.28 SOM: 6.74



Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 02	[30]	1-D -> Radial	SigmaZero T1 EtaZero T2	Euclidian	3.5	No β γ - -	.02	0	3	70	30	40	40	Yes	k-means: 69.66 SOM: 4.49



Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 03	[50]	1-D -> Radial	SigmaZero T1 EtaZero T2	Euclidian	4	No β γ - -	.02	0	3	70	30	40	40	Yes	k-means: 91.01 SOM: 1.12



تا اینجای کار ما در همهی حالات معماري تک بعدی برای لایهی کوهون، نوع همسایگی را قابل تشخیص و از نوع همسایگی شعاعی و در فاز مرتب شدن و فاز همگرائی تعداد ایپکها را به ترتیب ۴۰ و ۴۰ در نظر گرفته و همانطور که مشهود است در همهی حالات سعی کردهایم با در نظر گرفتن یک معیار خوشبندی مناسب(همان حد آستانه) به تعداد خوشههای مطلوب دادههای ورودی یعنی همان سه عدد دست پیدا کنیم که در همگی آزمایشها این مهم میسر شده است.

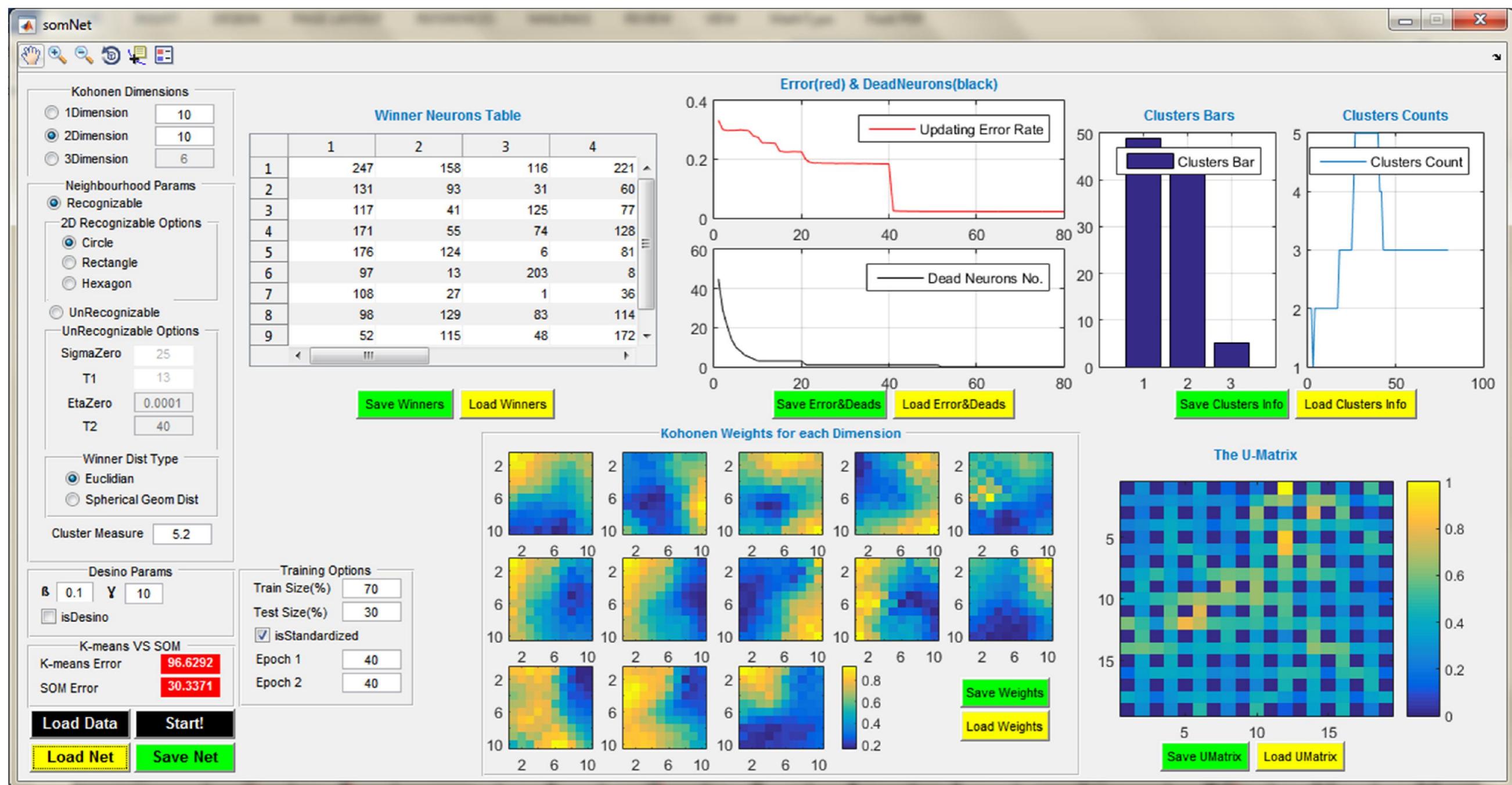
از نمودار شمار خوشهها (Clusters Count) پیداست که تعداد خوشهها در طول ایپکهای آموزشی تا حدی نوسان دارد اما در نهايیت به یک مقدار ثابت و مطلوب میل می کند و علت اين امر را نيز می توان در جابجاشدن نورونها و تغيير فاصله بين آنها در طول آموزش شبکه جستجو کرد. تعداد نورونهای مرده در پایان آموزش برابر صفر و میزان خطای بهروزشدن وزن نورونها (همان میزان جابجائی آنها) در هر سه حالت بعد از حدود ۴۰ ایپک (اتمام فاز مرتب شدن نورونها) به یک مقدار خيلي کم میل کرده و ثابت میماند و اين مسئله نشان می دهد که شبکه به یک حالت نسبتا پايدار رسیده است.

همچنین از روی نمودار U-Matrix نيز می توان تا حدودی مرز بين خوشههای تولید شده را تشخیص داد که اين مرز همان خطوط روشن بين نورونها می باشد که هرچه اين خط روشن تر باشد نمایانگر فاصله زياد بين دو نورون و هرچه تيره تر باشد نشان از مجاورت بيشتر بين نورونها دارد و به عبارتی نورونهای با اين آرایش در يك خوشه قرار دارند (لازم به ذكر است که در اينجا در حالت يك بعدی، U-Matrix به دليل استفاده ازتابع imagesc در متلب با عرض زياد و شبيه به حالت دو بعدی ترسيم شده است).

در نهايیت پیداست که در بين معماري های يك بعدی آزمایش شده بهترین نتيجه مربوط به حالت آرایش ۵۰ نورون در لایهی کوهون می باشد که خطای تست خوشبندی شبکهی کوهون در آن برابر ۱,۱۲ و به عبارتی زير ۵ درصد و خطای تست خوشبندی حاصل از روش K-Means نيز بالاي ۹۰ درصد می باشد. البته باید قيد کرد که خطای حاصل از روش K-Means همیشه تصادفی بوده و اينجانب موفق نشدم تا خطای مربوطه را با روش هایي مانند استفاده از تابع rng در متلب به مقدار ثابتی رسانده و يا حداقل از نوسان آن بعد از هر بار تست جلوگيري نمایم.

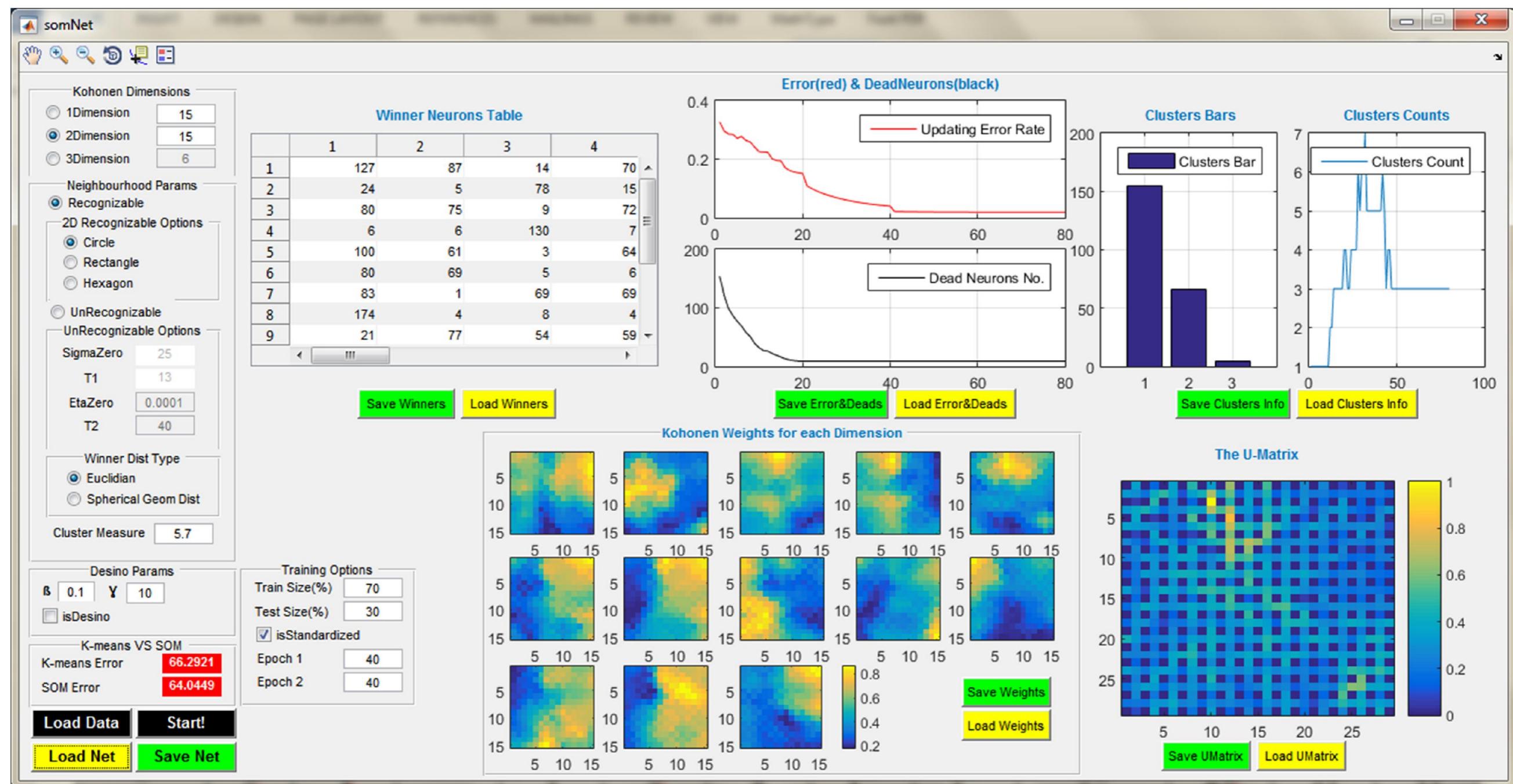
آزمایش چهارم (حالت دوبعدی - همسایگی دایره‌ای)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 04	[10, 10]	2-D -> Circular	SigmaZero T1 EtaZero T2	Euclidian	5.2	No β γ - -	.02	0	3	70	30	40	40	Yes	k-means: 96.62 SOM: 30.34



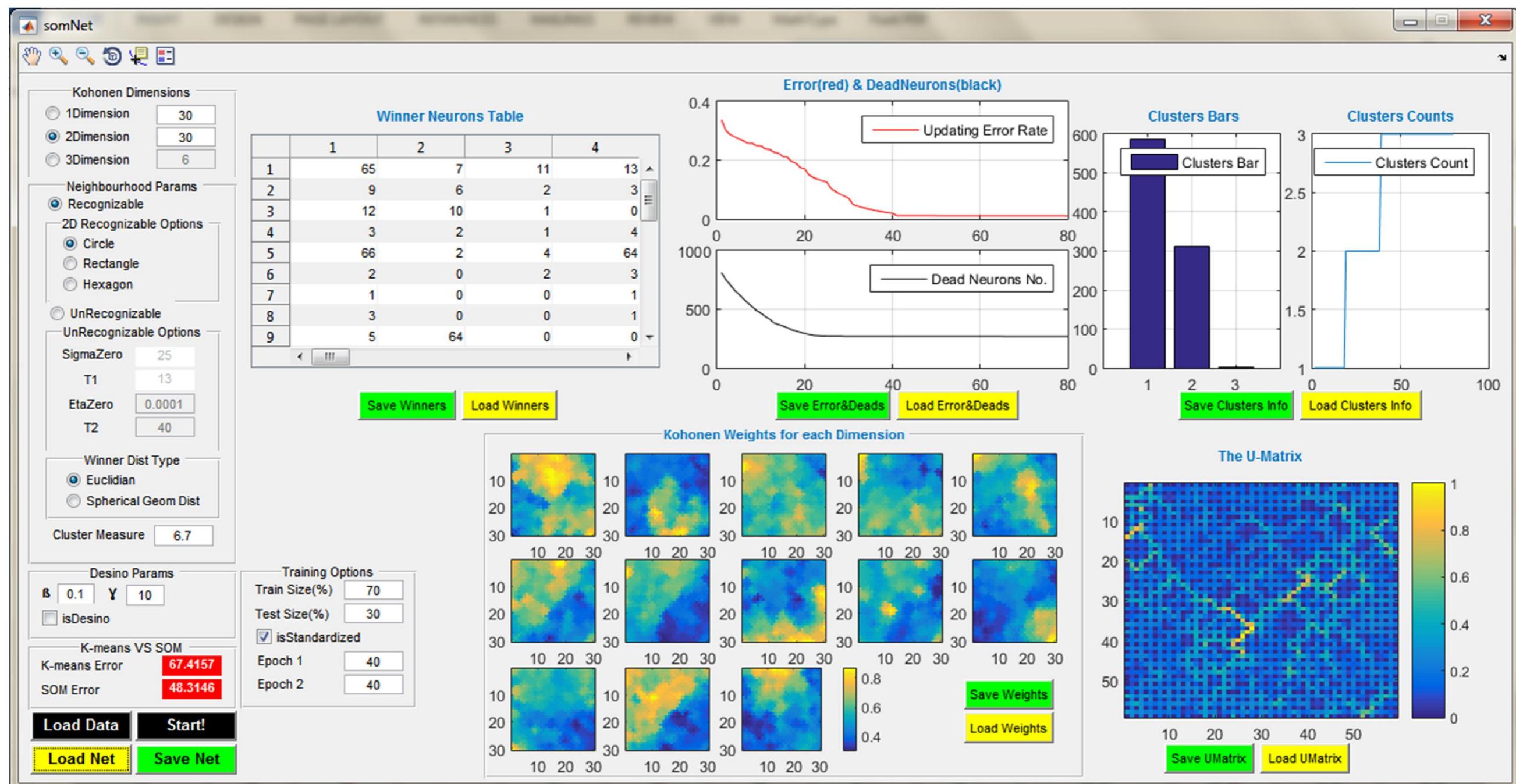
آزمایش پنجم (حالت دو بعدی - همسایگی دایره ای)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 05	[15, 15]	2-D -> Circular	SigmaZero T1 EtaZero T2	Euclidian	5.7	No β γ - -	.02	9	3	70	30	40	40	Yes	k-means: 66.29 SOM: 64.04



آزمایش ششم (حالت دو بعدی - همسایگی دایره ای)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 06	[30, 30]	2-D -> Circular	SigmaZero T1 EtaZero T2	Euclidian	6.7	No β γ - -	.01	267	3	70	30	40	40	Yes	k-means: 67.42 SOM: 48.31



تا اینجای کار نیز ما در همهٔ حالات معماری دو بعدی آزمایش شده برای لایهٔ کوهون، نوع همسایگی را قابل تشخیص و از نوع همسایگی دایره‌ای و تعداد ایپک‌های فازهای مرتب شدن و همگرائی را مانند قبل در نظر گرفته و همانطور که مشهود است باز هم در همهٔ حالات سعی کرد هایم با در نظر گرفتن یک حد آستانه‌ی مناسب به تعداد خوشه‌های مطلوب داده‌های ورودی یعنی همان سه عدد دست پیدا کنیم که در این آزمایش‌ها به آن دست یافته‌ایم. از نمودار شمار خوشه‌ها (Clusters Count) پیداست که به دلیل زیاد شدن نورون‌ها به دلیل افزایش ابعاد و نیز عدم یکنواخت بودن میزان برنده شدن نورون‌ها فاصله‌ی بین نورون‌ها متغیر بوده و در نتیجه تعداد خوشه‌ها نوسان دارد، اما در نهایت به یک مقدار ثابت و مطلوب میل می‌نماید.

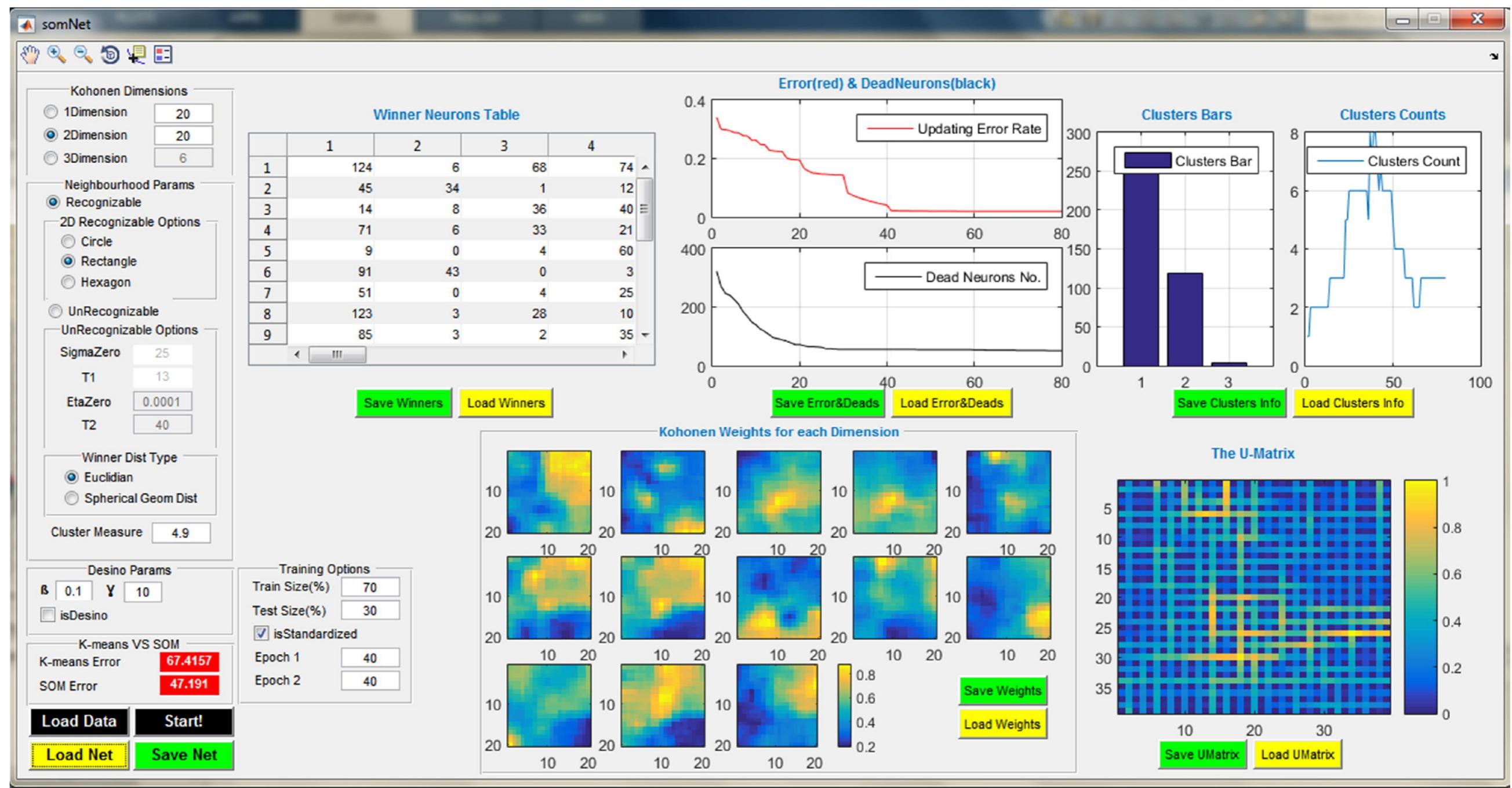
تعداد نورون‌های مرده در پایان آموزش برای معماری [10, 10] برابر صفر و برای سایر معماری‌های پیاده شده در حالت همسایگی دایره‌ای با افزایش تعداد نورون‌ها زیاد شده و در طول روند آموزش ثابت می‌ماند و علت آن نیز می‌تواند شاید عدم یکنواخت بودن توزیع داده‌های ورودی باشد. میزان خطای به روز شدن وزن نورون‌ها (همان میزان جابجایی آن‌ها) در هر سه حالت بعد از حدود ۴۰ ایپک (اتمام فاز مرتب شدن نورون‌ها) مانند حالت یک بعدی به یک مقدار خیلی کم میل کرده و ثابت می‌ماند و این مسئله نشان می‌دهد که شبکه به یک حالت نسبتاً پایدار رسیده است.

همچنین مانند حالت یک بعدی از روی نمودار U-Matrix نیز می‌توان تا حدودی مرز بین خوشه‌های تولید شده را تشخیص داد که این مرز همان خطوط روشن بین نورون‌ها می‌باشد که هرچه این خط روشن‌تر باشد نمایانگر فاصله زیاد بین دو نورون و هرچه تیره‌تر باشد نشان از مجاورت بیشتر بین نورون‌ها دارد و به عبارتی نورون‌های با این آرایش در یک خوشه قرار دارند.

در نهایت پیداست که در بین معماری‌های دو بعدی آزمایش شده با همسایگی از نوع دایره‌ای بهترین نتیجه مربوط به حالت آرایش [10, 10] می‌باشد که خطای تست خوشه‌بندی شبکه‌ی کوهون در آن برابر ۳۰,۳۴ و خطای تست خوشه‌بندی حاصل از روش K-Means نیز بالای ۹۰ درصد می‌باشد. همانطور که پیداست این شبکه نسبت به شبکه برگزیده‌ی یک بعدی مانند نتیجه‌ی بهتری را ارائه نمی‌دهد اما باز هم تا حدی قابل قبول است.

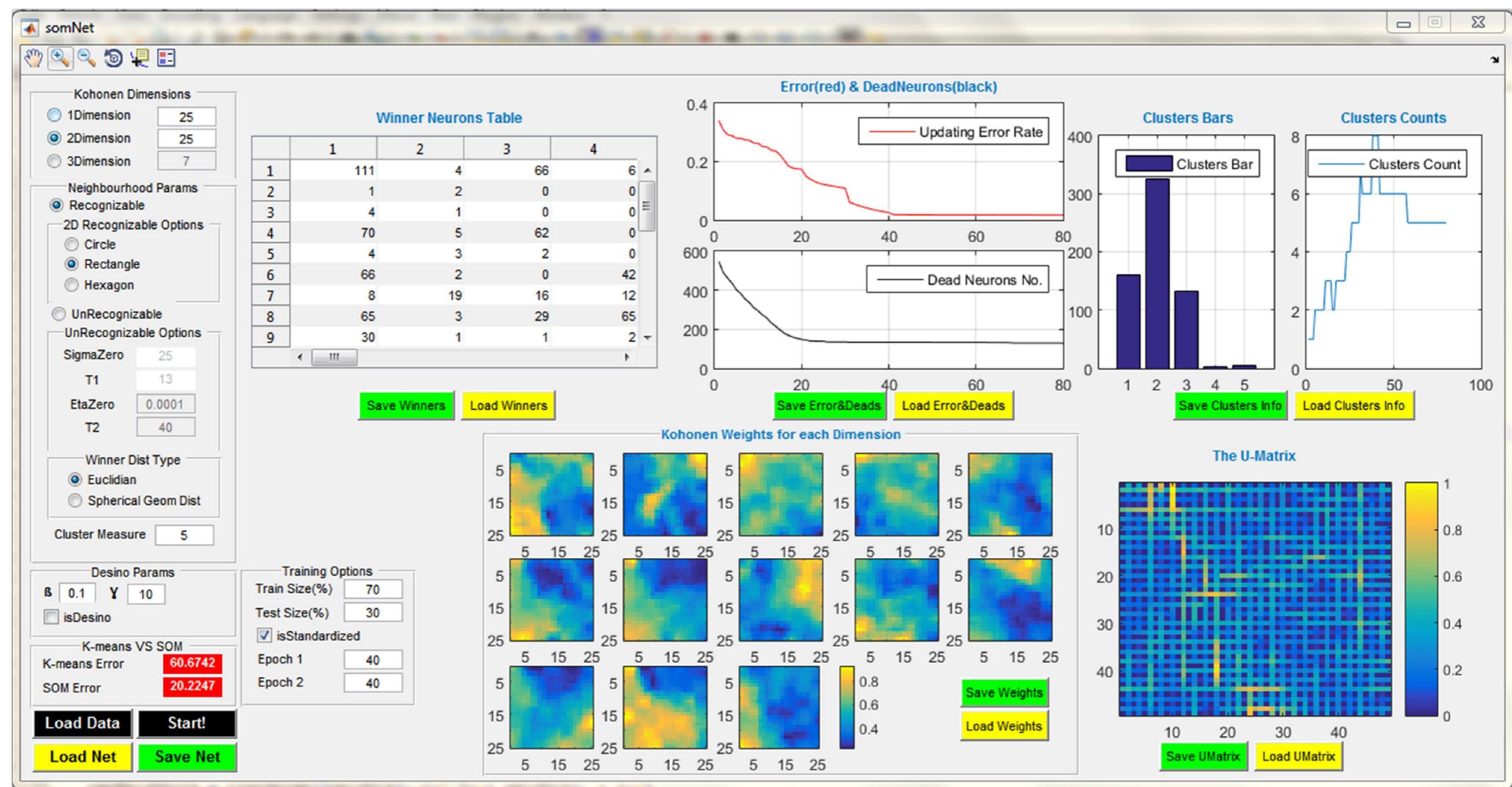
آزمایش هفتم (حالت دو بعدی - همسایگی مستطیلی)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 07	[20, 20]	2-D -> Rectangular	SigmaZero T1 EtaZero T2	Euclidian	4.9	No β γ - -	.02	51	3	70	30	40	40	Yes	k-means: 67.42 SOM: 47.19



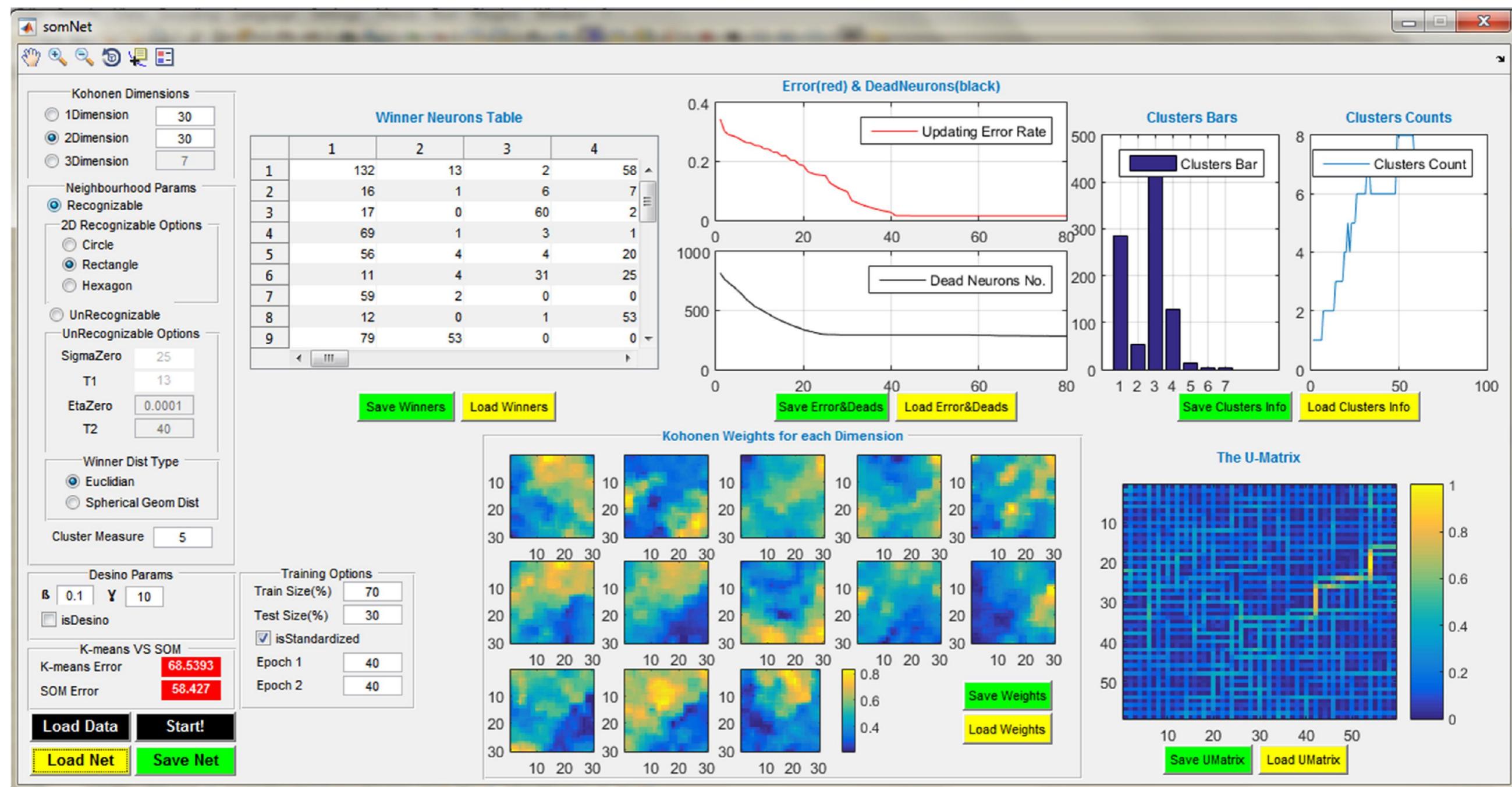
آزمایش هشتم (حالت دو بعدی - همسایگی مستطیلی)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 08	[25, 25]	2-D -> Rectangular	SigmaZero T1 EtaZero T2	Euclidian	5	No β γ	.02	129	5	70	30	40	40	Yes	k-means: 60.67 SOM: 20.22



آزمایش نهم (حالت دو بعدی - همسایگی مستطیلی)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors
Train 09	[30, 30]	2-D -> Rectangular	SigmaZero T1 EtaZero T2	Euclidian	5	No β γ	.01 283	7	70	30	40	40	Yes	k-means: 68.54 SOM: 58.43



تا اینجای کار نیز ما در حالات معماري دو بعدی آزمایش شده برای لایه‌ی کوهون، نوع همسایگی را قابل تشخیص و از نوع همسایگی مستطیلی و تعداد ایپک‌های فازهای مرتب شدن و همگرائی را مانند قبل در نظر گرفته و همانطور که مشهود است باز هم در همه‌ی حالات سعی کرده‌ایم با در نظر گرفتن یک حد آستانه‌ی مناسب به تعداد خوشه‌های مطلوب داده‌های ورودی یعنی همان سه عدد دست پیدا کنیم که در اینجا تنها در مورد آزمایش هفتم با معماری [20, 20] موفق به این کار شده و در مورد آزمایش‌های هشتم و نهم به دلیل زیاد شدن تعداد نورون‌ها و نیز توضیحاتی که در مورد معماری‌های پیشین قید شد حصول تنها سه خوشه بسیار دشوار گردید.

از نمودار شمار خوشه‌ها (Clusters Count) پیداست که به دلیل زیاد شدن نورون‌ها به دلیل افزایش ابعاد و نیز عدم یکنواخت بودن میزان برنده شدن نورون‌ها فاصله‌ی بین نورون‌ها متغیر بوده و در نتیجه تعداد خوشه‌ها نوسان دارد، اما در نهایت به یک مقدار ثابت و مطلوب میل می‌نماید. تعداد نورون‌های مرده برای معماری‌های پیاده شده با این شیوه‌ی همسایگی در طول آموزش هیچ وقت به صفر نرسیده و بعد از چند ایپک به یک مقدار ثابت غیر صفر میل می‌کند و علت آن نیز می‌تواند شاید عدم یکنواخت بودن توزیع داده‌های ورودی و نیز وجود افزونگی در تعداد نورون‌های لایه‌ی کوهون باشد.

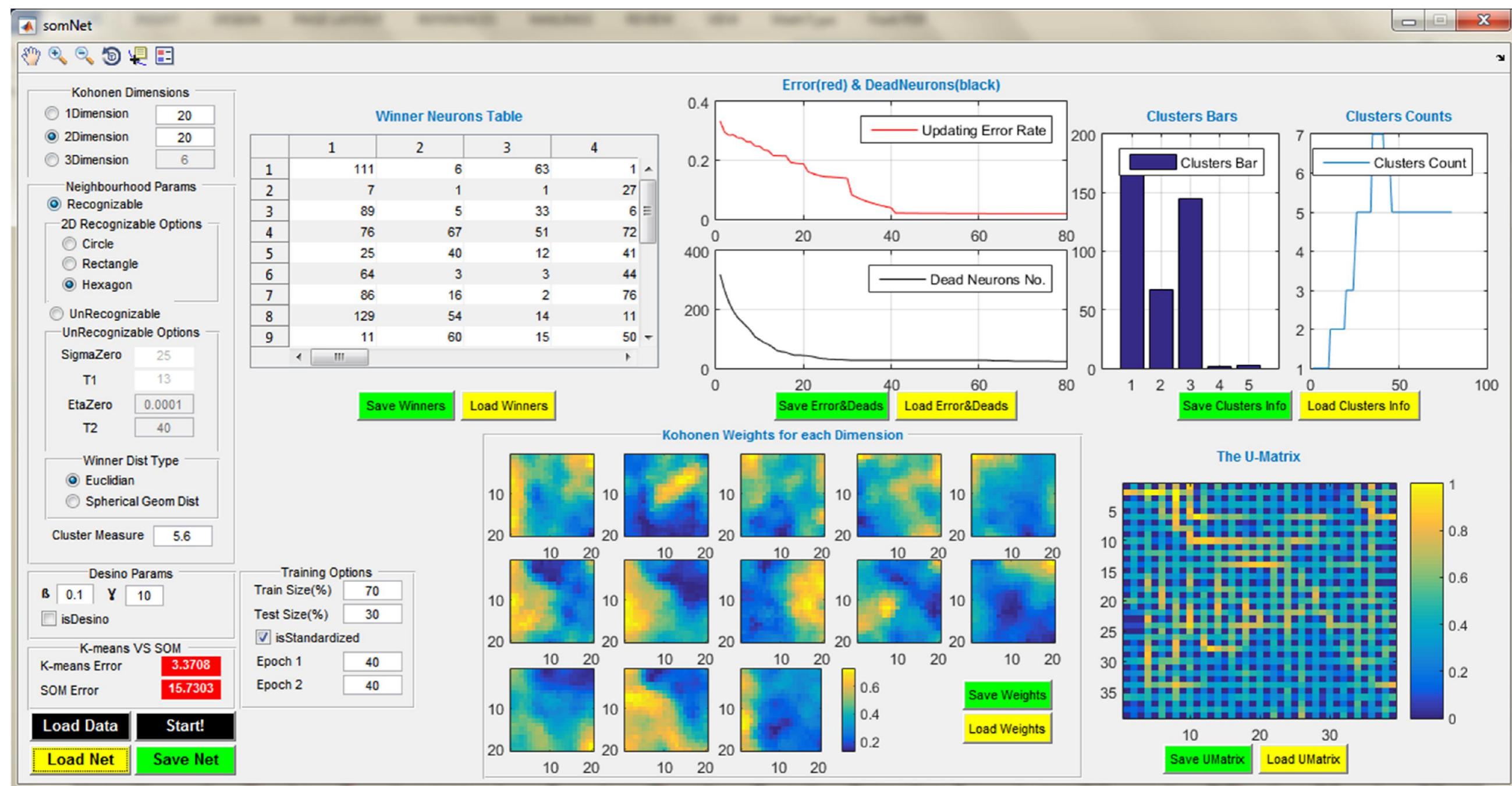
میزان خطای به روز شدن وزن نورون‌ها (همان میزان جابجائی آن‌ها) در هر سه حالت بعد از حدود ۴۰ ایپک (اتمام فاز مرتب شدن نورون‌ها) مانند حالات پیشین به یک مقدار خیلی کم میل کرده و ثابت می‌ماند و این مسئله نشان می‌دهد که شبکه به یک حالت نسبتاً پایدار رسیده است.

همچنین مانند حالات قبل از روی نمودار U-Matrix نیز می‌توان تا حدودی مرز بین خوشه‌های تولید شده را تشخیص داد. در نهایت پیداست که در بین معماری‌های دو بعدی آزمایش شده با همسایگی از نوع مستطیلی بهترین نتیجه مربوط به حالت آرایش [25, 25] می‌باشد که خطای تست خوشبندی شبکه‌ی کوهون در آن برابر ۲۰,۲۲ و خطای تست خوشبندی حاصل از روش K-Means نیز حدود ۶۰ درصد می‌باشد.

علت نتیجه‌ی بهتری که در این حالت همسایگی نسبت به همسایگی دایره‌ای حاصل شده نیز می‌تواند پوشش بیشتر همسایگی مستطیلی نسبت به همسایگی دایره‌ای با توجه به مستطیلی بودن آرایش نورون‌های لایه‌ی کوهون باشد. همانطور که پیداست این شبکه نیز نسبت به شبکه برگزیده‌ی یک بعدی ما نتیجه‌ی را ارائه نمی‌دهد اما باز هم تا حدی قابل قبول است.

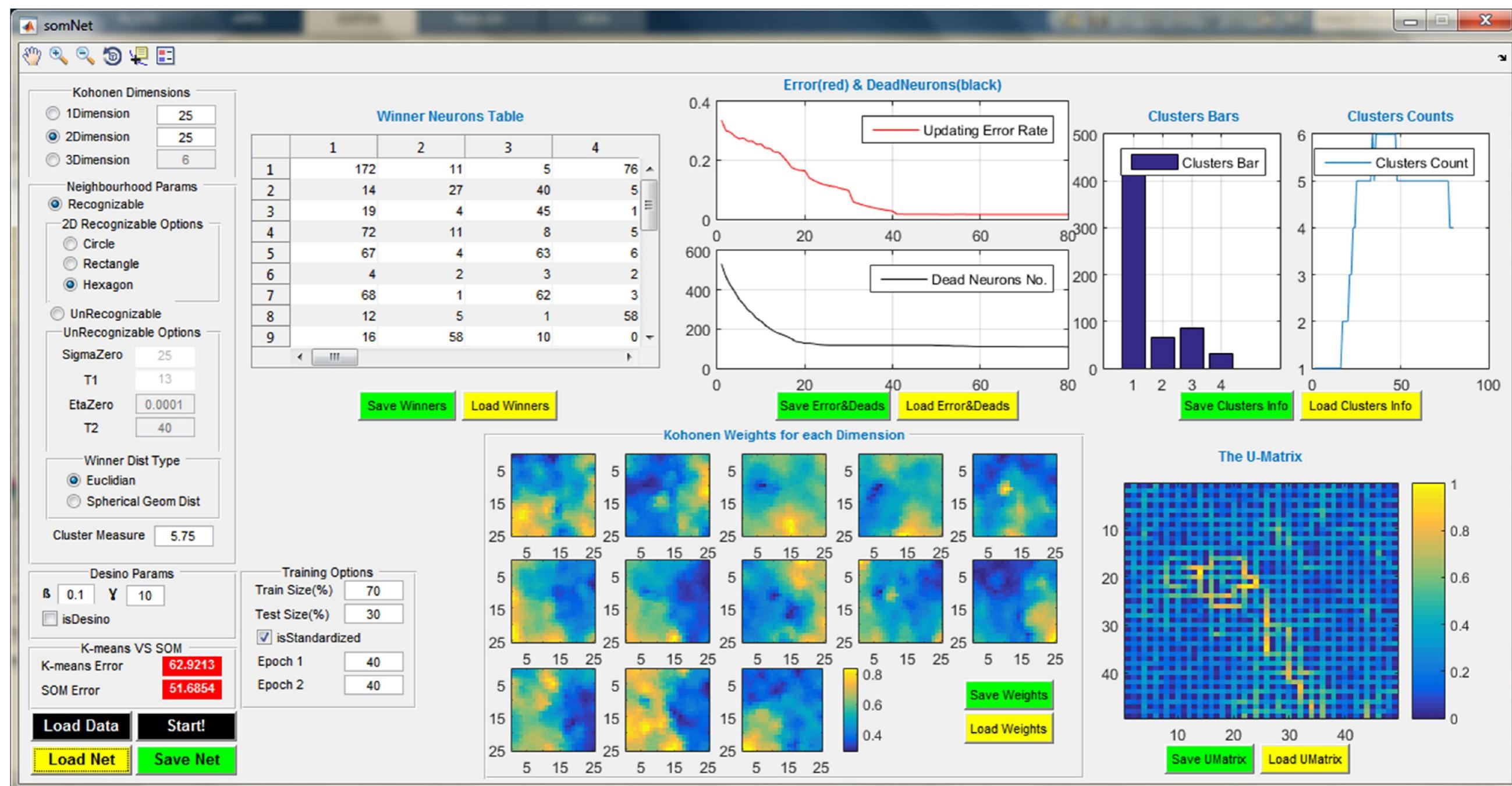
آزمایش دهم (حالت دوبعدی - همسایگی شش ضلعی)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 10	[20, 20]	2-D -> Hexagonal	SigmaZero T1 EtaZero T2 - - .1 -	Euclidian	5.6 β γ - -	No	.02	23	5	70	30	40	40	Yes	k-means: 3.37 SOM: 15.73



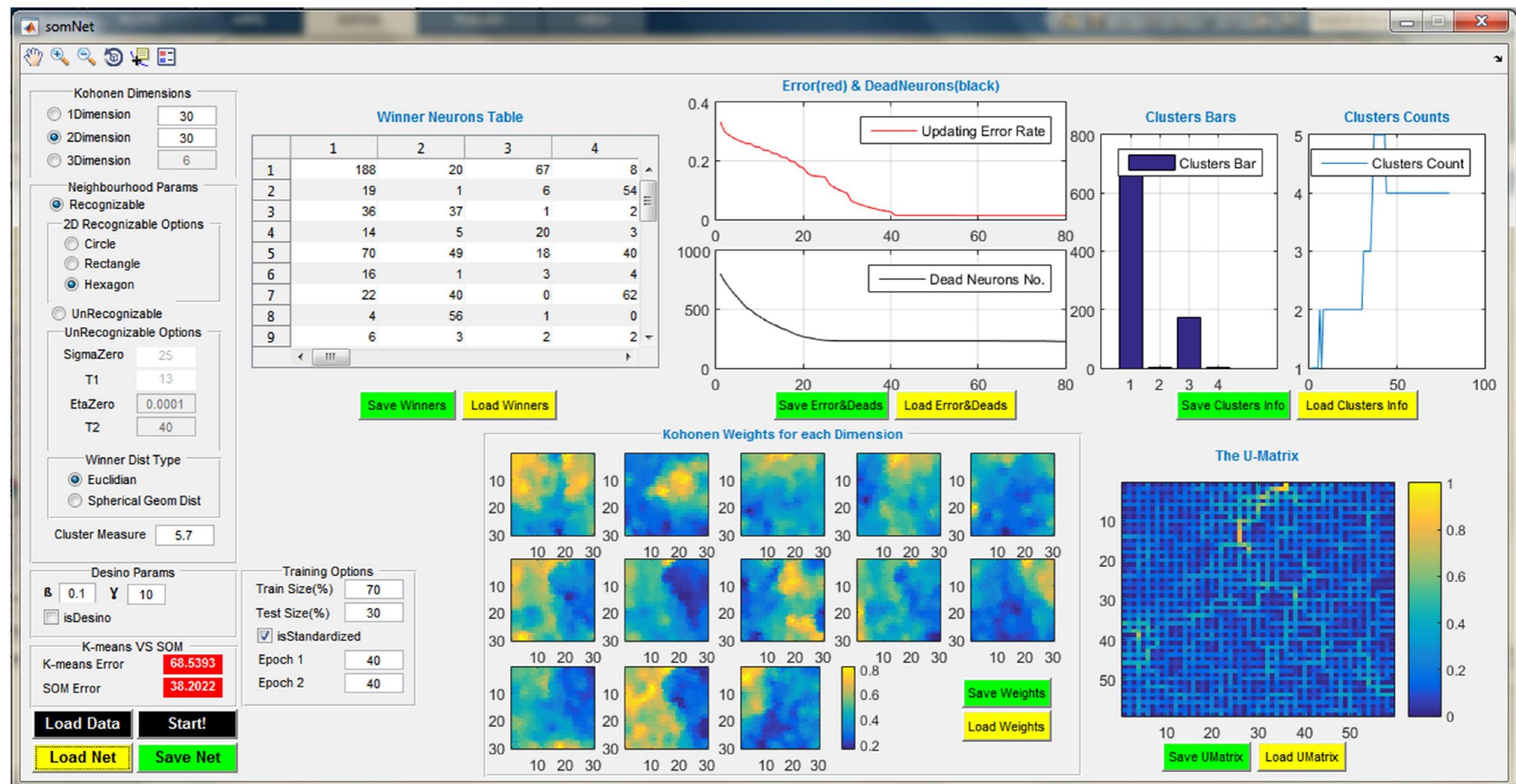
آزمایش یازدهم (حالت دو بعدی - همسایگی شش ضلعی)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 11	[25, 25]	2-D -> Hexagonal	SigmaZero T1 EtaZero T2	Euclidian	5.75	No β γ - -	.02	108	4	70	30	40	40	Yes	k-means: 62.92 SOM: 51.69



آزمایش دوازدهم (حالت دو بعدی - همسایگی شش ضلعی)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 12	[30, 30]	2-D -> Hexagonal	SigmaZero T1 EtaZero T2	Euclidian	5.7	No β γ - -	.01	227	4	70	30	40	40	Yes	k-means: 68.54 SOM: 38.20



تا اینجای کار نیز ما در حالات معماري دو بعدی آزمایش شده برای لایه‌ی کوهون، نوع همسایگی را قابل تشخیص و از نوع همسایگی شش ضلعی و تعداد ایپک‌های فازهای مرتب شدن و همگرائی را مانند قبل در نظر گرفته و همانطور که مشهود است باز هم در همه‌ی حالات سعی کرده‌ایم با در نظر گرفتن یک حد آستانه‌ی مناسب به تعداد خوشه‌های مطلوب داده‌های ورودی یعنی همان سه عدد دست پیدا کنیم که در اینجا متأسفانه در هیچ‌کدام از موارد موفق به این کار نشده و علت آن نیز مانند آنچه در مورد معماري‌های پیشین گفته شد وجود افزونگی در نورون‌های لایه‌ی کوهون و نیز محتمل‌بودن عدم وجود یکنواختی در توزیع داده‌های ورودی می‌باشد.

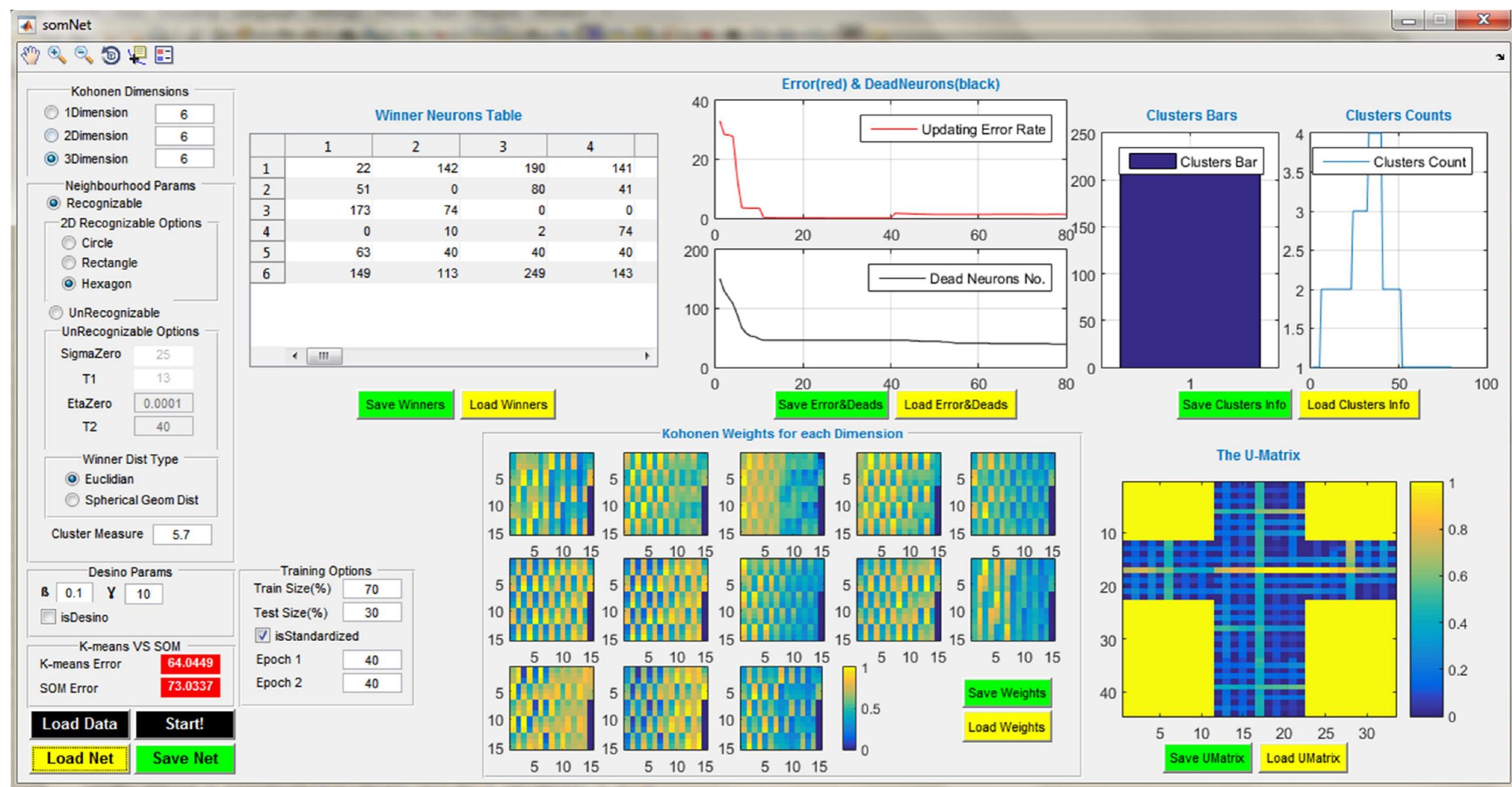
از نمودار شمار خوشه‌ها (Clusters Count) پیداست که به دلیل زیاد شدن نورون‌ها به دلیل افزایش ابعاد و نیز عدم یکنواخت بودن میزان برنده شدن نورون‌ها، فاصله‌ی بین نورون‌ها متغیر بوده و در نتیجه تعداد خوشه‌ها نوسان دارد، اما در نهایت به یک مقدار ثابت و مطلوب میل می‌نماید. تعداد نورون‌های مرده برای معماري‌های پیاده شده با این شیوه‌ی همسایگی در طول آموزش هیچ‌وقت به صفر نرسیده و بعد از چند ایپک به یک مقدار ثابت غیر صفر میل می‌کند و علت آن نیز مانند آنچه در مورد معماري‌های پیشین قید شد می‌باشد.

میزان خطای به روز شدن وزن نورون‌ها (همان میزان جابجائی آن‌ها) در هر سه حالت بعد از حدود ۴۰ ایپک (اتمام فاز مرتب شدن نورون‌ها) مانند حالات پیشین به یک مقدار خیلی کم میل کرده و ثابت می‌ماند و این مسئله نشان می‌دهد که شبکه به یک حالت نسبتاً پایدار رسیده است. همچنین مانند حالات قبل از روی نمودار U-Matrix نیز می‌توان تا حدودی مرز بین خوشه‌های تولید شده را تشخیص داد.

در نهایت پیداست که در بین معماري‌های دو بعدی آزمایش شده با همسایگی از نوع شش ضلعی بهترین نتیجه مربوط به حالت آرایش [20, 20] می‌باشد که خطای تست خوشه‌بندی شبکه‌ی کوهون در آن برابر ۱۵.۷۳ و خطای تست خوشه‌بندی حاصل از روش K-Means نیز در اینجا به اتفاق مقدار کمی می‌باشد. همانطور که پیداست این شبکه نیز نسبت به شبکه برگزیده‌ی یک بعدی ما نتیجه‌ی بهتری را ارائه نمی‌دهد اما باز هم تا حدی قابل قبول است.

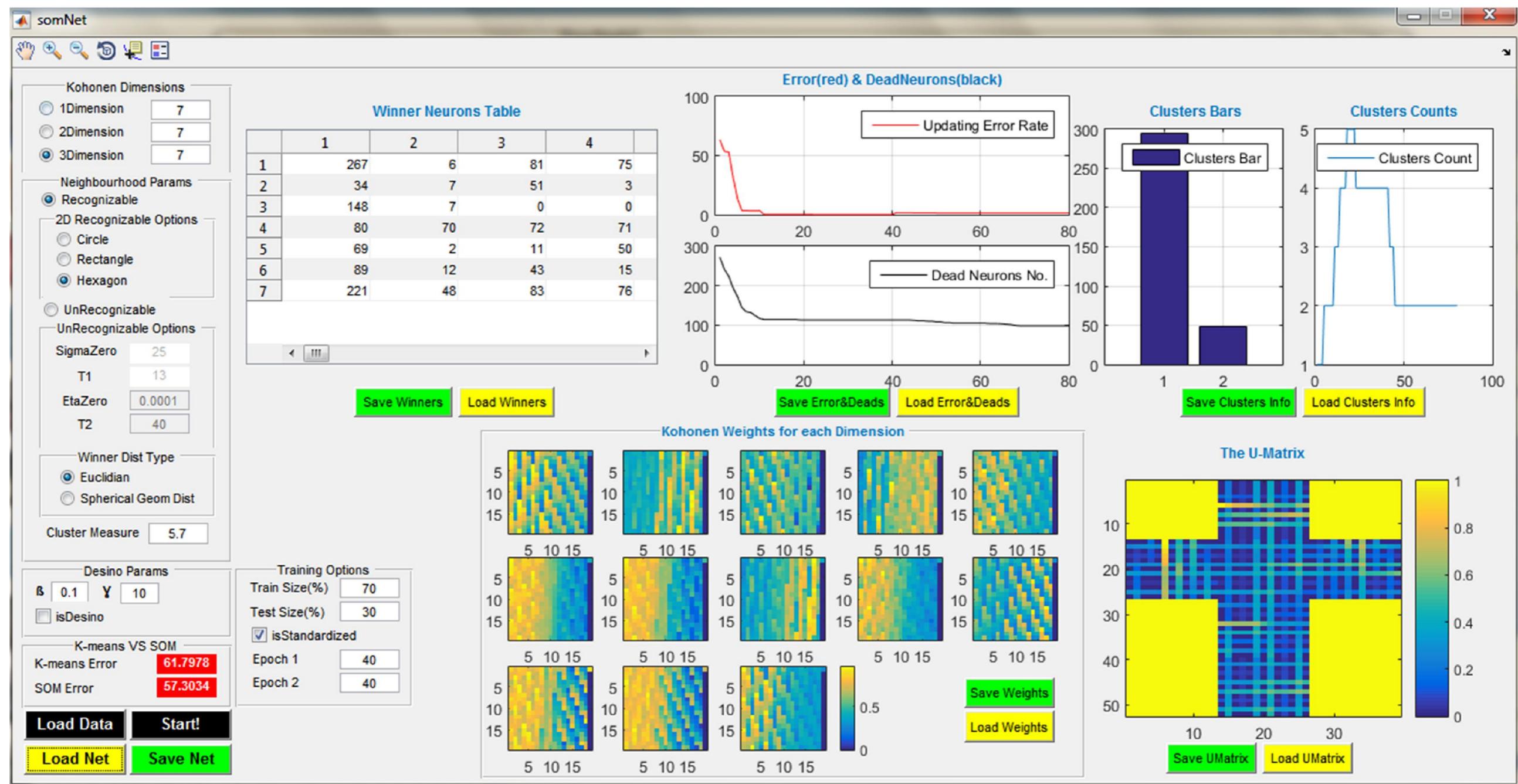
آزمایش سیزدهم (حالت سه بعدی - همسایگی مکعبی)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors
Train 13	[6, 6, 6]	3-D -> Cubic SigmaZero T1 EtaZero T2 - - .1 -	Euclidian	5.7 β γ - -	No	1.26	39	1	70	30	40	40	Yes	k-means: 64.04 SOM: 73.03



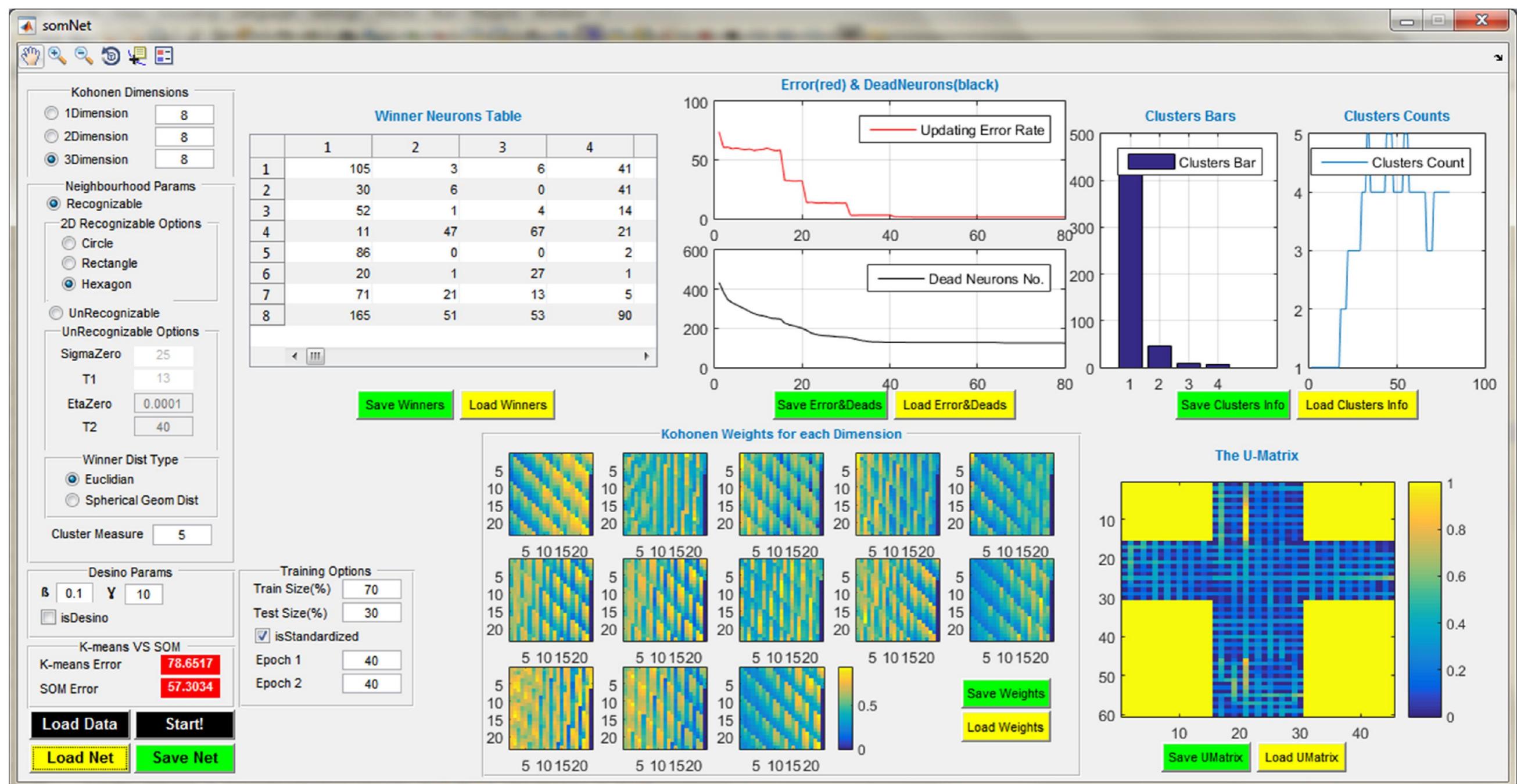
آزمایش چهاردهم (حالت سه بعدی - همسایگی مکعبی)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 14	[7, 7, 7]	3-D -> Cubic	SigmaZero T1 EtaZero T2	Euclidian	5.7	No β γ - -	1.33	98	2	70	30	40	40	Yes	k-means: 61.80 SOM: 57.30



آزمایش پانزدهم (حالت سه بعدی - همسایگی مکعبی)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 15	[8, 8, 8]	3-D -> Cubic	SigmaZero T1 EtaZero T2	Euclidian	5	No β γ	1.32	124	4	70	30	40	40	Yes	k-means: 78.65 SOM: 57.30



تا اینجای کار نیز برای حالات معماري سه بعدی آزمایش شده برای لایه‌ی کوهون، نوع همسایگی را قابل تشخیص و از نوع همسایگی مکعبی و تعداد ایپک‌های فازهای مرتب شدن و همگرائی را مانند قبل در نظر گرفته و همانطور که مشهود است باز هم در همه‌ی حالات سعی کرده‌ایم با در نظر گرفتن یک حد آستانه‌ی مناسب به تعداد خوشه‌های مطلوب داده‌های ورودی یعنی همان سه عدد دست پیدا کنیم که در اینجا متأسفانه در هیچ‌کدام از موارد موفق به این کار نشده و علت آن نیز مانند آنچه در مورد معماري‌های پیشین گفته شد وجود افزونگی در نورون‌های لایه‌ی کوهون و نیز محتمل‌بودن عدم وجود یکنواختی در توزیع داده‌های ورودی می‌باشد.

از نمودار شمار خوشه‌ها (Clusters Count) پیداست که تعداد خوشه‌ها نوسان داشته و در نهایت به یک مقدار ثابت و مطلوب میل می‌نماید. تعداد نورون‌های مرده برای معماري‌های پیاده‌شده با این شیوه‌ی همسایگی در طول آموزش هیچ‌وقت به صفر نرسیده و بعد از چند ایپک به یک مقدار ثابت غیر صفر میل می‌کند و علت آن نیز مانند آنچه در مورد معماري‌های پیشین قید شد می‌باشد.

میزان خطای به روزشدن وزن نورون‌ها (همان میزان جابجایی آن‌ها) در هر سه حالت بعد از حدود ۴۰ ایپک (اتمام فاز مرتب شدن نورون‌ها) مانند حالات پیشین به یک مقدار خیلی کم میل کرده و ثابت می‌ماند و این مسئله نشان می‌دهد که شبکه به یک حالت نسبتاً پایدار رسیده است.

همچنین مانند حالات قبل از روی نمودار U-Matrix نیز که در اینجا تنها شش وجه مکعب نورون‌ها را ترسیم کرده‌ایم می‌توان تا حدودی مرز بین خوشه‌های تولید شده را تشخیص داد.

در نهایت پیداست که در بین معماري‌های سه بعدی آزمایش شده با همسایگی از نوع مکعبی نتایج تست شبکه در هر دو حالت روش‌های کوهون و البته در همگی حالات حدوداً مشابه یکدیگر می‌باشند چرا که در اینجا ما به دلیل طولانی شدن زمان آزمایش شبکه به دلیل زیاد شدن تعداد نورون‌ها در حالت آرایش سه بعدی از به کار بردن نورون‌های زیاد پرهیز کرده و در نتیجه نتایج مشابه حاصل می‌شوند و البته از نمودارهای مربوطه پیداست که بعد از گذشت چند ایپک وضعیت نمودارها چندان تغییری ندارند و در نتیجه شاید بتوان گفت که افزایش نورون‌ها چندان نتیجه‌ی بهتری را حاصل ننماید. همانطور که پیداست این شبکه نیز نسبت به شبکه برگزیده‌ی یک بعدی ما نتیجه‌ی بهتری را ارائه نمی‌دهد و در نتیجه ما نیز در ادامه این گزارش برای سایر موارد خواسته شده در تعریف پروژه از همین معماري بهره خواهیم برداشت.

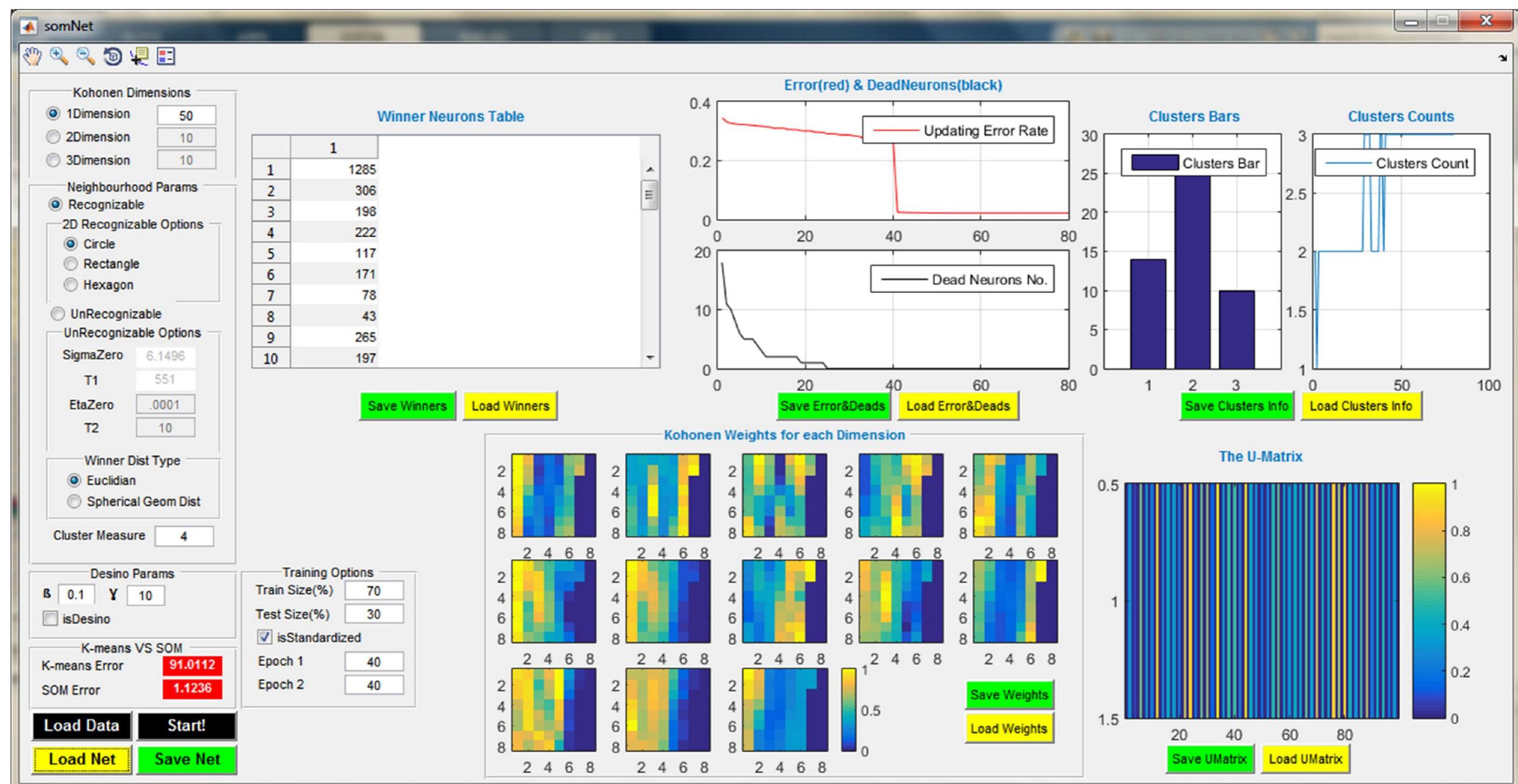
نتیجه گیری:

با توجه به آزمایش‌های انجام‌شده تاکنون و البته مطالب بیان‌شده برای هر کدام از معماری‌ها که بعضاً دارای حشو و زیاده‌گوئی نیز بودند! می‌توان چنین نتیجه گرفت که با افزایش ابعاد لایه‌ی کوهون و به دنبال آن زیادشدن تعداد نورون‌ها، نه تنها شاهد نتیجه‌ی بهتری در تعداد خوش‌ها و نیز نتیجه تست خواهیم بود بلکه تعداد نورون‌های مرده نیز با این کار افزایش می‌یابد و علت این امر را نیز می‌توان در محتمل‌بودن عدم یکنواختی توزیع داده‌های ورودی دانست که سبب ایجاد نواحی ایزوله شده و در نتیجه تعدادی از نورون‌ها در این نواحی گیر کرده و هیچ‌گاه وزنشان به روز نشده و به عبارتی به نورون مرده تبدیل می‌شوند. در نهایت در ادامه‌ی این گزارش همانطور که پیش از این قید شد ما همان معماری تک‌بعدی با ۵۰ نورون لایه‌ی کوهون را بر سایر معماری‌ها ترجیح داده و سایر مقایسه‌ها را با همین معماری انجام خواهیم داد.

۲. بررسی تأثیر نوع فاصله قابل تشخیص و غیر قابل تشخیص در مرز بین خوشه‌های تولیدشده در U-Matrix

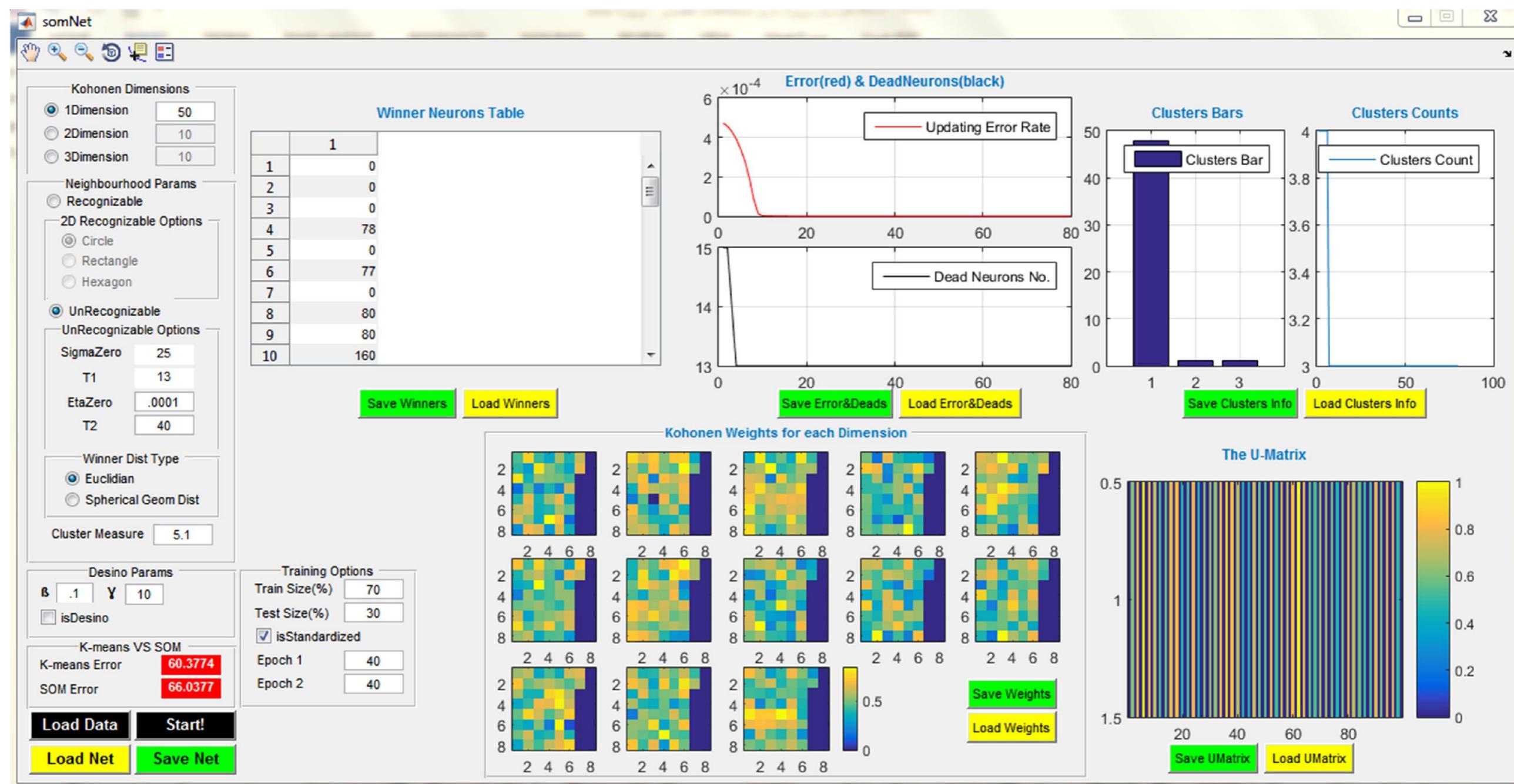
آزمایش شانزدهم (فاصله قابل تشخیص)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors
Train 16	[50]	1-D -> Radial	SigmaZero T1 EtaZero T2	Euclidian	4	No β γ	.02 0	3 70	30 40	40 40	Yes	k-means: 91.01 SOM: 1.12		



آزمایش هفدهم (فاصله غیرقابل تشخیص)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 17	[50]	1-D -> Unrecognizable	SigmaZero T1 EtaZero T2 25 13 .0001 40	Euclidian	5.1	No β γ - -	8.529e-14	13	3	70	30	40	40	Yes	k-means: 60.38 SOM: 66.04



نتیجه‌گیری:

در این آزمایش‌ها سعی داریم تفاوت وضعیت شبکه در دو حالت همسایگی قابل تشخیص و غیر قابل تشخیص را مورد بررسی قرار دهیم. در حالت همسایگی غیر قابل تشخیص (آزمایش هفدهم) که تشابه زیادی به عملکرد نورون‌های طبیعی داشته و در آن یک واحد فعال بر روی واحدهای همسایه‌ی نزدیک‌تر بیشتر تأثیرگذار است تا واحدهای دورتر، پس از تشخیص نورون برنده به عبارتی نه تنها وزن این نورون بلکه سایر نورون‌ها نیز باشد و ضعف متفاوت به روزرسانی می‌شوند و به عبارتی انتظار این است که وزن همگی نورون‌ها به نحوی اصلاح گردد. اما باز هم شاید به دلیل وجود عدم یکنواختی در توزیع داده‌های ورودی و البته زیادبودن نورون‌های لایه‌ی کوهونن برخی نورون‌ها فرصت به روزشدن پیدا نکرده و حتی با زیاد نمودن تعداد ایپک‌های آموزشی همچنان تا آخر آموزش مرده باقی می‌مانند. نکته‌ی دیگر اینکه در این روش همسایگی ما پارامترهای T_2 که همان تعداد ایپک‌های فاز مرتب‌شدن آموزش بوده و البته η_0 که همان مقدار اولیه برای نرخ یادگیری می‌باشد را به صورت دستی مقدار اولیه داده و سپس پارامتر T_1 را با توجه به فرمول $T_1 = \log \sigma_0 / T_2$ و پارامتر σ_0 را نیز با توجه به شعاع نقشه‌ی نورون‌های لایه‌ی کوهونن که برابر بیشینه‌ی نصف هر کدام از ابعاد این لایه می‌باشد، مقدار می‌دهیم و در فرمول‌های مربوط به به روزرسانی وزن نورون‌ها در این حالت همسایگی به کار می‌بریم. در اینجا ما برخلاف حالات همسایگی با مرز قابل تشخیص مقدار اولیه نرخ یادگیری را برای جلوگیری از همگراشدن همگی نورون‌ها به یک نقطه برابر یک مقدار خیلی کمتر یعنی 0.0001 در نظر می‌گیریم که البته این مقدار در طول فاز مرتب‌شدن طبق فرمول گاوسین کاهش یافته و سپس در طول فاز همگرائی ثابت می‌ماند.

در مورد نرخ خطای نورون‌ها نیز می‌بینیم بعد از حدود ۱۰ ایپک و در همان ایپک‌های فاز مرتب‌شدن، نرخ خطای صفر رسیده و تا آخر آموزش بلا تغییر باقی می‌ماند و به عبارتی شبکه به حالت پایداری رسیده و در فاز همگرائی نیز شاهد هیچ تغییری نخواهیم بود.

تعداد خوش‌ها نیز با توجه به معیار خوش‌بندی در نظر گرفته شده بعد از تعدادی نوسانات به همان سه عدد خوش میل کرده و ثابت می‌ماند که با اینکه مسئله مطلوب ماست اما همچنان نرخ خطای شبکه‌ی کوهونن بالا بوده و در برابر بهترین معماری فعلی ما با فاصله‌ی از نوع قابل تشخیص توان ایستادگی نداشته و در ادامه نیز ما همان فاصله‌ی از نوع قابل تشخیص را ترجیح خواهیم داد.

۳. بررسی نتایج خوشه‌بندی حاصل از شبکه کوهون و خوشه‌بندی K-Means

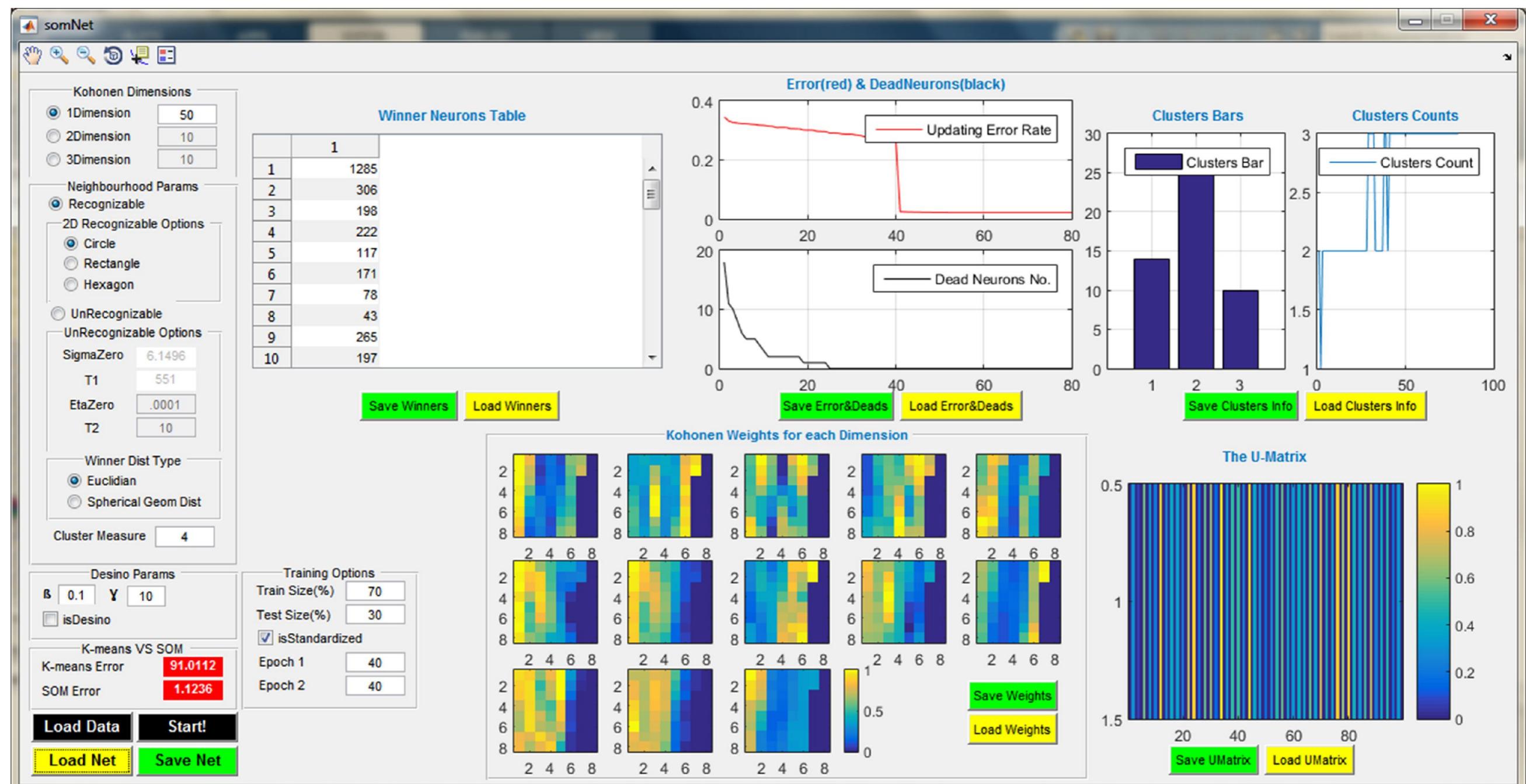
در تمامی آزمایش‌های پیشین همان‌طور که مشاهده گردید ما بعد از هر آزمایش با تعداد نصف داده‌های ورودی، شبکه را با هر دو روش کوهون و K-Means تست نمودیم. باید گفت که نتایج حاصله برای شبکه‌ی کوهون با توجه به اینکه هر بار یک سری داده‌ی تصادفی از بین داده‌های ورودی برای آموزش به شبکه ارائه می‌شوند، تا حدودی تصادفی بوده و باید تلاش کرد تا با انجام آزمایش‌های متعدد و درنظر گرفتن معیار خوشه‌بندی یا همان threshold مناسب به یک نتیجه‌ی تست مطلوب دست یافت که در بسیاری موارد موفق به این مهم نشدیم اما اگر شبکه‌ی ذخیره‌شده را دوباره بارگذاری نموده و همین عمل تست را روی آن انجام دهیم خواهیم دید که نتایج تست تا حدود زیادی مشابه پیشین بوده و این نشان از کسب یک آرایش خوب برای نورون‌های شبکه‌ی کوهون می‌باشد. در مورد تست با روش K-Means نیز همان‌طور که قید گردید از آن‌جا که مراکز خوشه‌ها توسط الگوریتم K-Means هر بار به صورت تصادفی تولید می‌شوند در نتیجه نتایج این روش نیز همیشه تصادفی می‌باشند و البته اینجانب در MATLAB با یک سری داده‌ی ثابت روش مربوطه آزمایش نمودم و هر بار نتیجه‌ی تست متفاوتی مشاهده گردید.

تا اینجا نیز معماری برگزیده‌ی ما همچنان پیروز میدان بوده و در موارد خواسته‌شده بعدی در تعریف پروژه از معماری نامبرده بهره خواهیم برد.

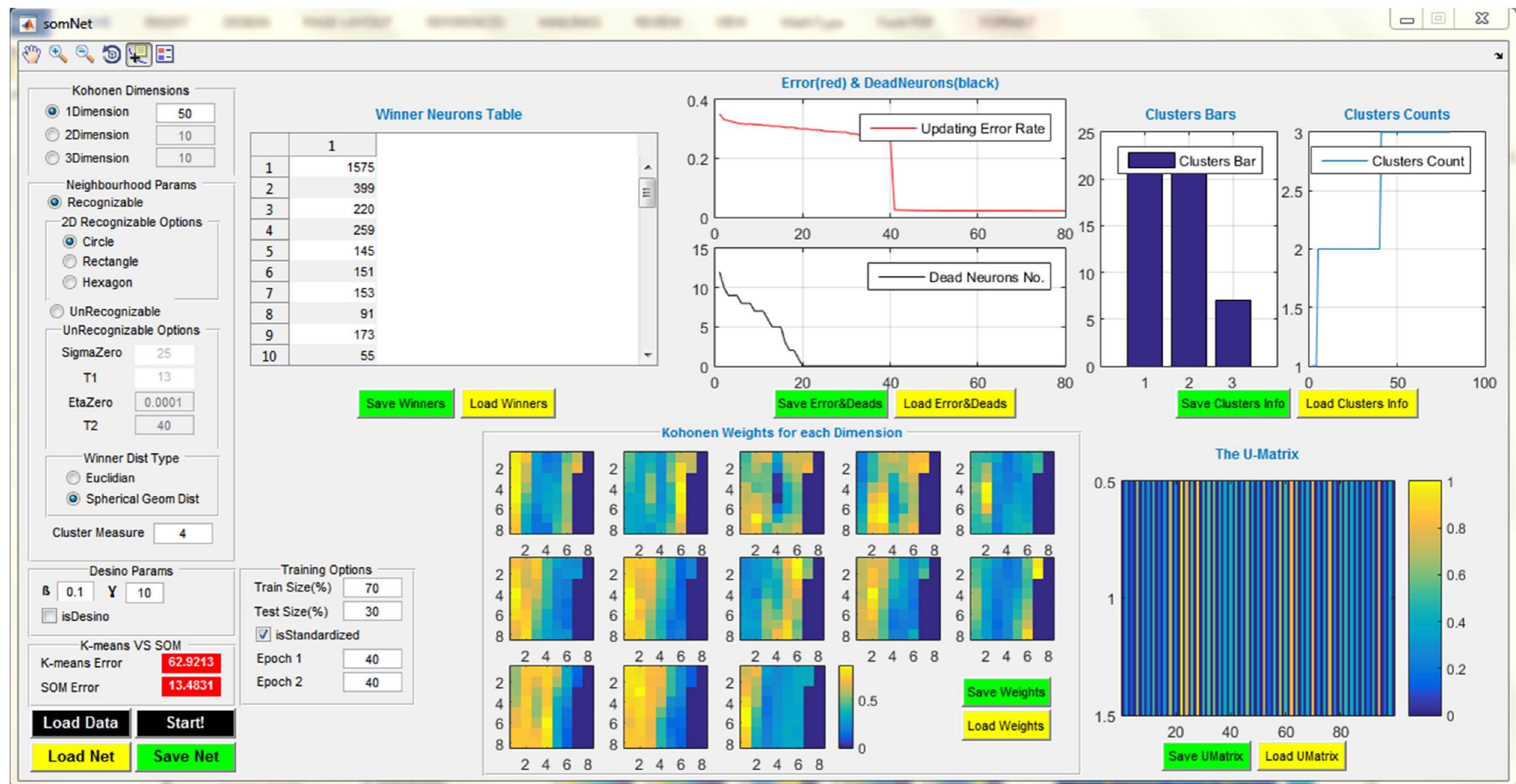
۴. بررسی تأثیر نوع فاصله‌ی مورد استفاده برای تعیین نورون برنده (فاصله‌ی اقلیدسی – فاصله‌ی کمانی‌کروی)

آزمایش هیجدهم (فاصله اقلیدسی)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 18	[50]	1-D -> Radial	SigmaZero T1 EtaZero T2	Euclidian	4	No β γ	.02	0	3	70	30	40	40	Yes	k-means: 91.01 SOM: 1.12



Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 19	[50]	1-D -> Unrecognizable	SigmaZero T1 EtaZero T2	Spherical Geometry Distance	4	No β γ -	.02	0	3	70	30	40	40	Yes	k-means: 62.92 SOM: 13.48



نتیجه گیری:

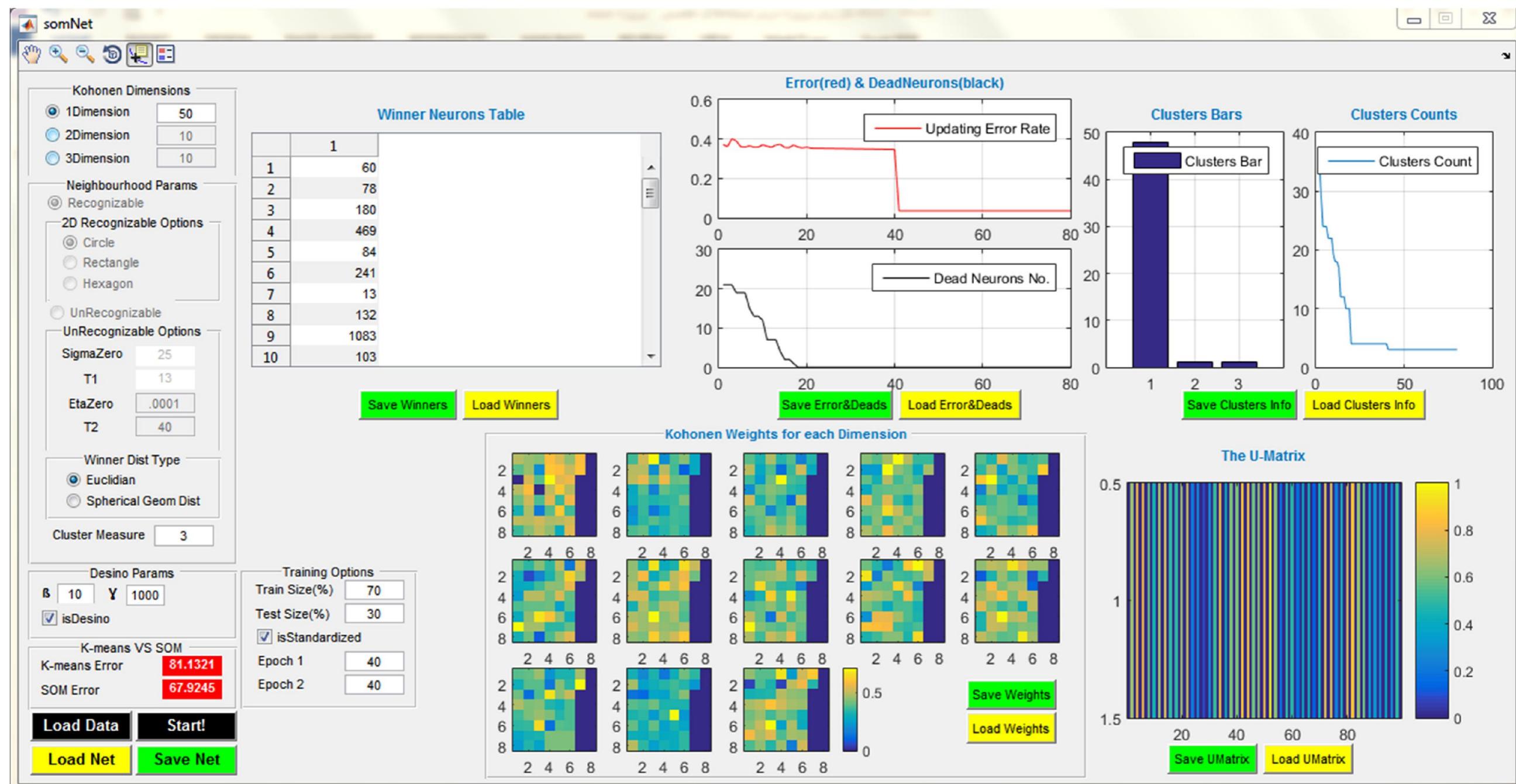
همان‌طور که پیداست بین استفاده از نوع فاصله‌ی اقلیدسی و فاصله‌ی کمانی‌کروی (Spherical Geometry Distance) چندان تفاوتی در تعداد خوش‌ها و البته نتایج تست مشاهده نمی‌گردد. شاید بتوان علت این امر را در تکیه‌ی اساس شبکه‌ی کوهون بر اصل مجاورت نورون‌ها به داده‌های ورودی و نه نوع مجاورت مربوطه دانست. چرا که بعد از پیداشدن نورون برنده، نورون‌های همسایه نیز بنا به نوع فاصله‌ی مربوطه تعیین شده و تفاوتی در اصل موضوع مشاهده نمی‌گردد.

همچنان معماری برگزیده حاصله تاکنون موفق عمل کرده و در ادامه‌ی گزارش مورد استفاده قرار خواهد گرفت.

۵. بررسی عملکرد راه حل دسینو در برابر سایر روش‌های مشتمل بر همسایگی

آزمایش بیستم (پیاده‌سازی راه حل دسینو)

Exp. No.	Kohonen Dimensions	Neighbourhood Type	Winner Distance Type	Cluster Measure	isDesino	Final Updating Error Rate	Final Dead Neurons No.	Clusters Count	Train Set Size(%)	Test Set Size(%)	Epoch1 (Arranging Phase)	Epoch2 (Converging Phase)	isStandardized	Errors	
Train 19	[50]	1-D > Unrecognizable	SigmaZero T1 EtaZero T2	Euclidian	3	No β 10 γ 1000	.04	0	3	70	30	40	40	Yes	k-means: 81.13 SOM: 67.92



نتیجه‌گیری:

همان‌طور که مشاهده می‌شود راه حل دسینو در اینجا با همان آرایش معماری برگزیده تاکنون یعنی ۵۰ نورون در لایه‌ی کوهون پیاده شده است. نکته اینجاست که با توجه به اینکه در روش دسینو بعد از تعیین خروجی همه‌ی نورون‌های لایه‌ی کوهون (صفر برای نورون بازنده و یک برای نورون برنده) و کسب مقادیر نسبت زمانی برندeshدن و بایاس برای همه‌ی نورون‌ها، نورونی به روزرسانی می‌گردد که تفاضل فاصله با داده‌ی ورودی و مقدار بایاس برای آن کمترین مقدار را داشته باشد، لذا در نهایت تنها وزن یک نورون در این روش به روز شده و همسایگی معنائی ندارد. پس در این شیوه ما همه‌ی پارامترهای روش‌های همسایگی از نوع قابل تشخیص و غیر قابل تشخیص را نادیده گرفته و در ظاهر گرافیکی غیرفعال می‌نماییم. مشاهده می‌گردد که تعداد نورون‌های مرده در این روش با درنظر گرفتن پارامترهای بهینه برای راه حل دسینو از جمله $\beta = 10$ و $\gamma = 1000$ بعد از تعداد ۱۸ ایپک آموزشی به صفر می‌رسد که این خود نشان از صحت عملکرد روش دسینو در جلوگیری از بیش از اندازه برندeshدن نورون‌ها داشته و نرخ خطای نیز بعد حدود ۴۰ ایپک آموزشی یعنی همان فاز مرتب‌شدن به مقدار اندکی میل کرده و همچنان تا پایان آموزش ثابت می‌ماند و این مطلب نشان از رسیدن شبکه به یک حالت پایدار دارد.

با وجود این‌که در این روش تعداد خوش‌های نهائی حاصله با در نظر گرفتن معیار خوش‌بندی مناسب برابر مقدار مطلوب می‌باشد اما نرخ خطای تست شبکه مقبول نبوده و در نتیجه در پایان این گزارش می‌توان گفت که بهترین نوع معماری در این پروژه همان معماری تک‌بعدی با ۵۰ نورون در لایه‌ی کوهون می‌باشد. چرا که هم نرخ خطای پایینی داشته و نیز تعداد خوش‌ها را درست تشخیص داده و همینطور خطای تست خوبی را ارائه می‌نماید.

در مورد راه حل دسینو و این‌که چرا نتیجه‌ی نهائی مطلوب نمی‌باشد می‌توان گفت که از آنجا که داده‌های آموزشی ورودی هر بار تصادفی از بین کل داده‌ها انتخاب می‌شوند، لذا هر بار نتیجه‌ی متفاوتی مشاهده می‌گردد و با انجام تعداد آزمایش‌های بیشتر نیز موفق به کسب نتیجه‌ی بهتری نگشتم.