Gastric Cancer   - Alex Matov

1 - > p< 0.05
1 *** -> p<0.01

Sep 25, 2013

Figure 2

Diffuse GC cell lines
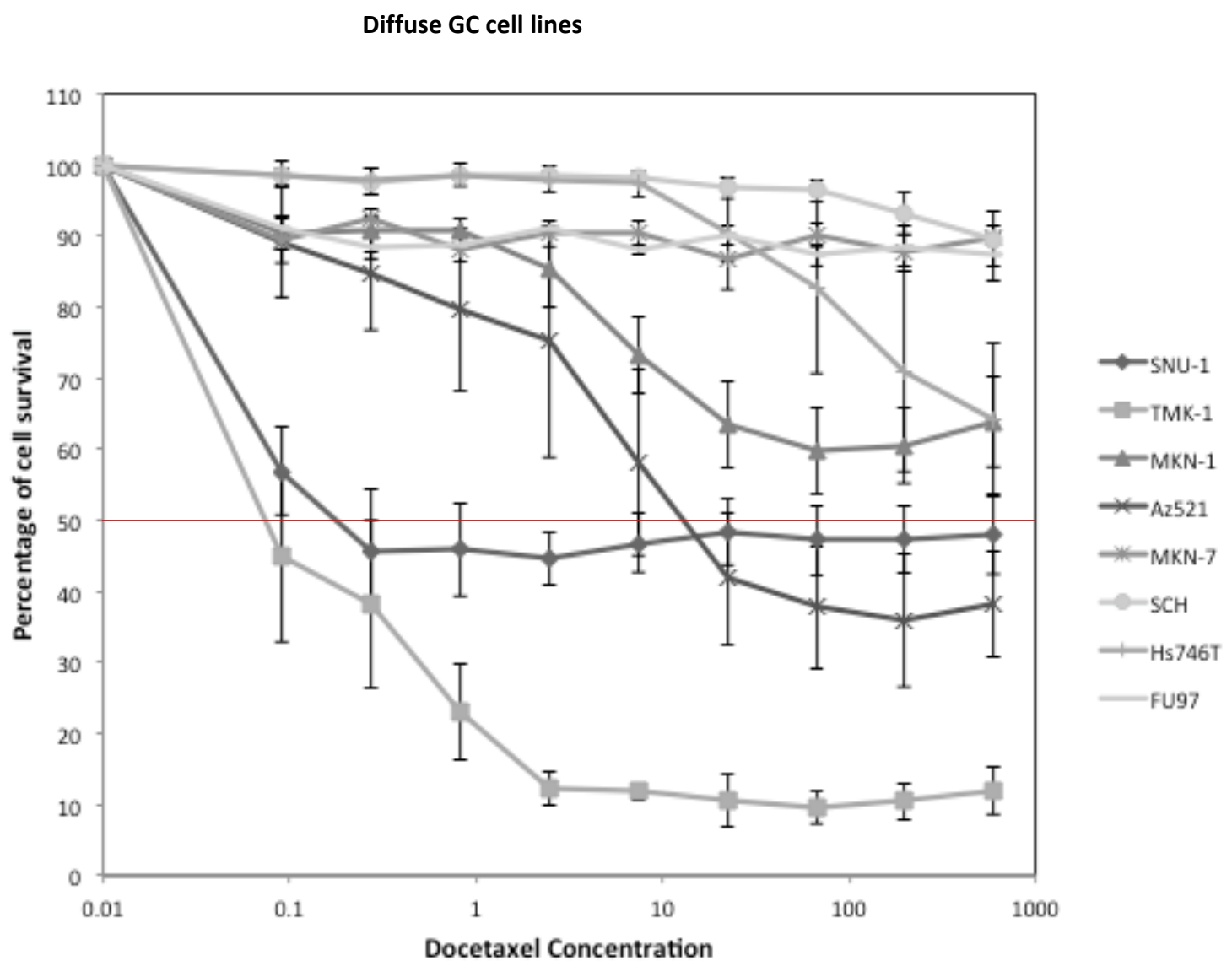
FEATURES

MORPH 1-7

1) object: number

2) object: EurlerNumber

3) object_size:average - **The average number of above-threshold pixels per object – captures information about the size of objects in the MT area of a cell**

4) object_size:variance - **The variance of the number of above-threshold pixels per object – quantifies the homogeneity of fluorescent objects in the image**

5) object_size:ratio - **The ratio of the size of the largest object to the smallest within the cell – assessing the distribution of florescent object sizes**

6) object_distance: average

7) object_distance: variance

EDGE 1-5

9) edges:area_fraction - **The fraction of the non-zero pixels in a cell that are along an edge** – distinguishes protein that localizes along the edges

10) edges:homogeneity - **Measure of edge intensity homogeneity** - captures homogeneity of edge gradients

11) edges:direction_maxmin_ratio **- Measure of edge direction homogeneity**

12) edges:direction_maxnextmax_ratio

13) edges:direction_difference - **Measure of edge direction difference - this feature distinguish MT patterns in which there are parallel edges**

SKELET 1-3

14) obj_skel_len - **The average length of the morphological skeleton of objects**

15) obj_skel_hull_area_ratio

18) obj_skel_branch_per_len - **The ratio of the number of branch points in skeleton to length of skeleton, averaged over all objects**: A point was defined as a branch point if 3 or more of its neighbors were contained within the skeleton.

CONVEX HULL 1-3

19) convex_hull: fraction of overlap -

20) convex_hull: shape_factor-

21) convex_hull: eccentricity -

Figure 2

	MKN-45	SNU-1	TMK-1	AZ-521	MKN-7	SCH	Hs746T
DMSO							
DTX 10nM							
DTX 100nM							
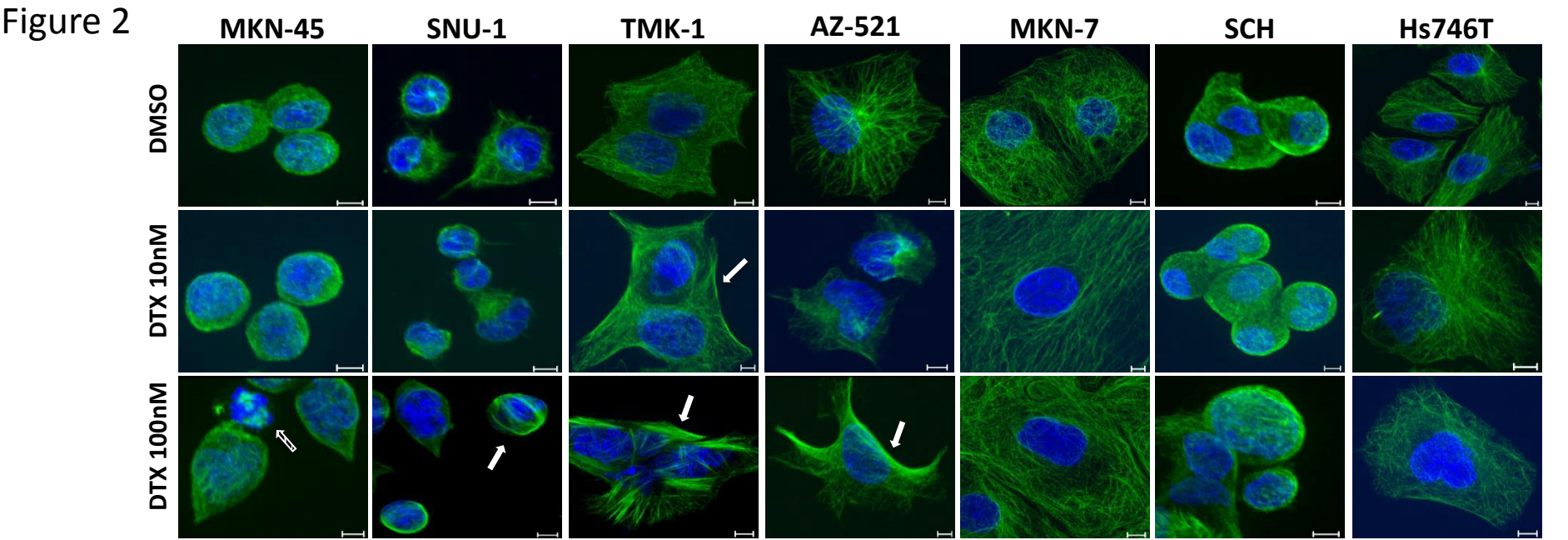
Feature	p<0.01 ***	4	11	5	1	3	2
Morph 1	repeat		1 ***				
Morph 2			1 ***				
Morph 3	1 *** -> p<0.01		1 ***				
Morph 4			1 ***				
Morph 5							
Morph 6					1 ***		
Morph 7			1 ***				
Edge 1		1 ***		1 ***			1 ***
Edge 2			1 ***				
Edge 3		1 ***	1 ***				1 ***
Edge 4		1 ***				1 ***	
Edge 5		1 ***					
Skelet 1			1 ***	1 ***			
Skelet 2				1 ***			
Skelet 3							
Convex Hull 1			1 ***	1 ***			
Convex Hull 2			1 ***	1 ***		1 ***	
Convex Hull 3			1 ***			1 ***	

Figure 2
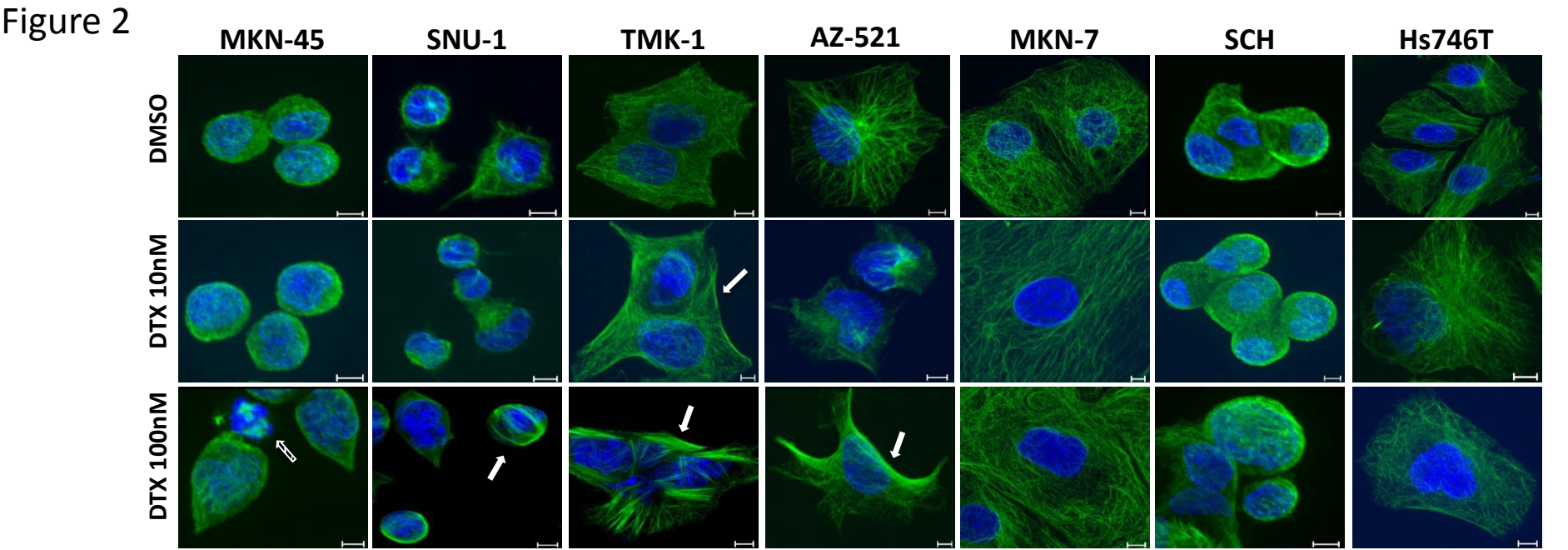
Feature	p<0.05	7	14	8	5	7	3
Morph 1	repeat		1 ***				
Morph 2		1	1 ***			1	
Morph 3		1	1 ***	1			
Morph 4			1 ***		1		
Morph 5			1	1	1		1
Morph 6			1		1 ***		
Morph 7			1 ***			1	
Edge 1		1 ***	1	1 ***		1	1 ***
Edge 2		1	1 ***				
Edge 3		1 ***	1 ***	1			1 ***
Edge 4		1 ***				1 ***	
Edge 5		1 ***					
Skelet 1			1 ***	1 ***	1		
Skelet 2				1 ***			
Skelet 3					1		
Convex Hull 1			1 ***	1 ***		1	
Convex Hull 2			1 ***	1 ***		1 ***	
Convex Hull 3			1 ***			1 ***	

1 - > p< 0.05

1*** - > p< 0.01

Table 2. Comparison of subcellular location feature sets. All features that measure length or area are calculated in pixels that are 0.23 μm square in the sample plane.

Feature description	SLF3	SLF7
Morphological features: Number of fluorescent objects in image, Euler number of image, average object size, variance of object size, ratio of largest to smallest object size, average object distance to cell center of fluorescence, variance of object distance to cell center, ratio of largest to smallest object distance to cell center	SLF1.1 through SLF1.8	SLF1.1 through SLF1.8
Edge-related features: Fraction of above-threshold pixels along edge, measure of edge gradient intensity homogeneity, measure of edge direction homogeneity 1, measure of edge direction homogeneity 2, measure of edge direction difference	SLF1.9 through SLF1.13	SLF7.9 through SLF7.13 (minor error corrections)
Convex hull features: Fraction of convex hull occupied by above-threshold pixels, roundess of convex hull, eccentricity of convex hull	SLF1.14 through SLF1.16	SLF1.14 through SLF1.16
Zernike moment features through order 12, calculated for a unit circle with radius equal to the average radius of the cell type being analyzed (150 pixels or 34.5 μm for HeLa)	SLF3.17 through SLF3.65	SLF3.17 through SLF3.65
Haralick texture features: angular second moment, contrast, correlation, sum of squares variation, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, info. measure of correlation 1, info. measure of correlation 2	SLF3.66 through SLF3.78	SLF7.66 through SLF7.78 (after downsampling to 1.15 μm/pixel and 256 gray levels)
Fraction of non-object fluorescence	–	SLF7.79
Skeleton features (see text)	–	SLF7.80 through SLF7.84

Feature Extraction
by Robert Murphy

Thresholding

- First type of feature is morphological
- Morphological features require some method for defining objects
- Most common approach is global thresholding

2D Features
Morphological Features

Description
The number of fluorescent objects in the image
The Euler number of the image
The average number of above-threshold pixels per object
The variance of the number of above-threshold pixels per object
The ratio of the size of the largest object to the smallest
The average object distance to the cellular center of fluorescence(COF)
The variance of object distances from the COF
The ratio of the largest to the smallest object to COF distance

2D Features
Morphological Features

DNA features (objects relative to DNA reference)

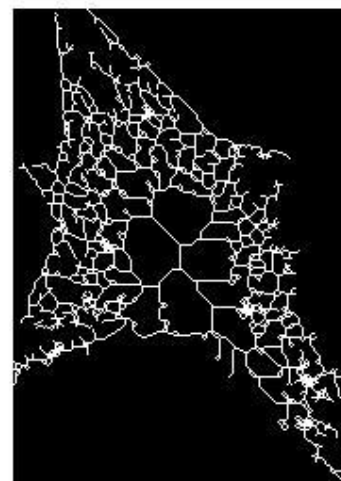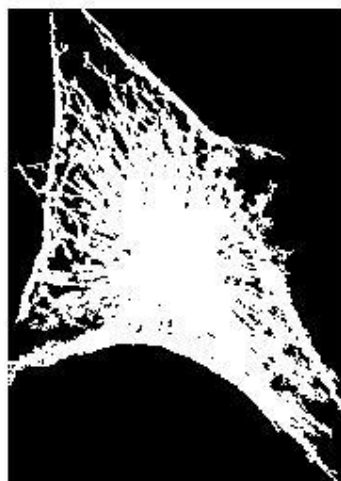| Description |
| --- |
| The average object distance from the COF of the DNA image |
| The variance of object distances from the DNA COF |
| The ratio of the largest to the smallest object to DNA COF distance |
| The distance between the protein COF and the DNA COF |
| The ratio of the area occupied by protein to that occupied by DNA |
| The fraction of the protein fluorescence that co-localizes with DNA |

2D Features
Morphological Features

Skeleton features

Description
The average length of the morphological skeleton of objects
The ratio of object skeleton length to the area of the convex hull of the skeleton, averaged over all objects
The fraction of object pixels contained within the skeleton
The fraction of object fluorescence contained within the skeleton
The ratio of the number of branch points in the skeleton to the length of skeleton

Illustration – Skeleton

2D Features
Edge Features

Edge features

| Description |
|---|
| The fraction of the non-zero pixels that are along an edge |
| Measure of edge gradient intensity homogeneity |
| Measure of edge direction homogeneity 1 |
| Measure of edge direction homogeneity 2 |
| Measure of edge direction difference |

2D Features
Hull Features

Convex hull (geometrical) features

The fraction of the convex hull area occupied by protein fluorescence
The roundness of the convex hull
The eccentricity of the convex hull

Zernike Moment Features

- Shape similarity of protein image to Zernike polynomials Z(n,l)
- 49 polynomials and 49 features
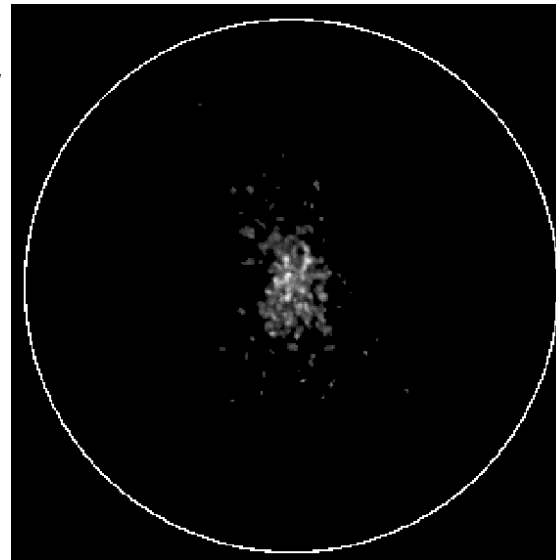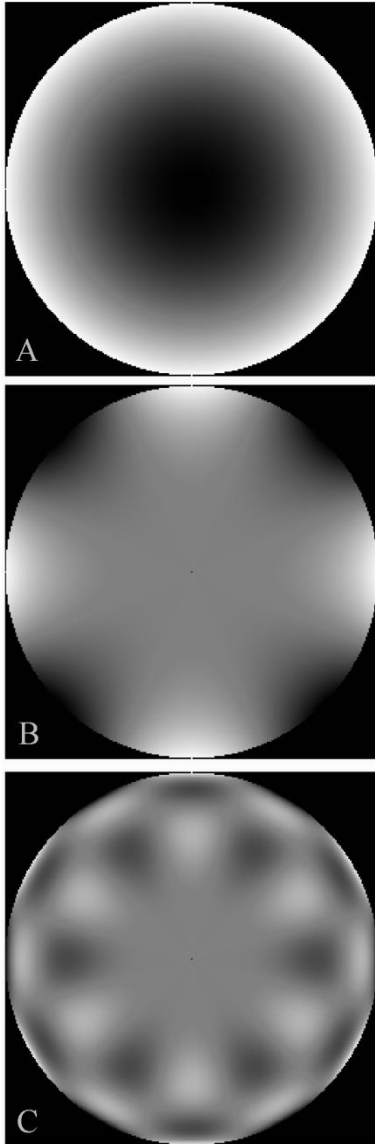
left: Zernike polynomials
A: Z(2,0)
B: Z(4,4)
C: Z(10,6)

right: lamp2 image
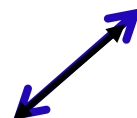
Haralick Texture Features

- Correlations of adjacent pixels in gray level images
- Start by calculating co-occurrence matrix P:

  N by N matrix, N=number of gray level.

  Element P(i,j) is the probability of pixels with value i being adjacent with pixels with value j

- Four directions in which a pixel can be adjacent

Co-occurrence Matrix
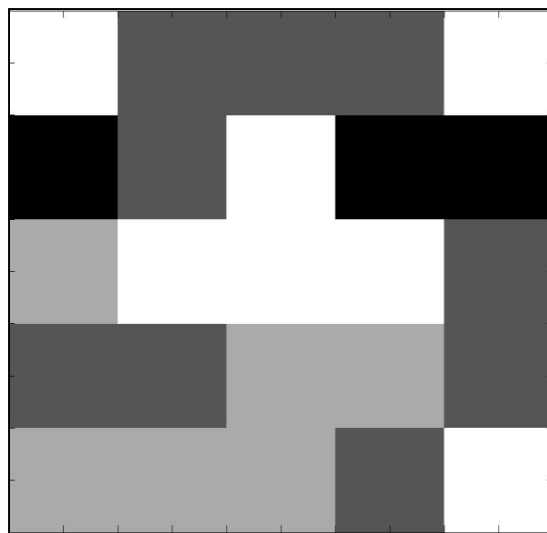
	1	2	3	4
1	0	2	1	3
2	2	4	4	4
3	1	4	2	2
4	2	3	2	2

	1	2	3	4
1	2	1	0	1
2	1	6	3	4
3	0	3	6	2
4	1	4	2	4

	1	2	3	4
1	0	1	0	3
2	1	4	3	3
3	0	3	4	1
4	3	3	1	2

	1	2	3	4
1	0	3	0	1
2	3	0	4	4
3	0	4	0	3
4	1	4	3	2