Appendix 1: Features recognized by Mazda

	meaning	*Texture features*
Gray histogram	the characters reflecting image uniformity	Mean (histogram's mean); Variance (histogram's variance); Skewness (histogram's skewness); Kurtosis (histogram's kurtosis); Perc.01% (1% percentile); Perc.10% (10% percentile); Perc.50% (50% percentile); Perc.90% (90% percentile); Perc.99% (99% percentile)
Gradient	a direction which change in grey scale intensity that representing the image intensity distribution	GrMean (absolute gradient mean); GrVariance (absolute gradient variance); GrSkewness (absolute gradient skewness); GrKurtosis (absolute gradient kurtosis); GrNonZeros (percentage of pixels with nonzero gradient)
Run-length matrix	calculated from angles of horizontal, vertical, 45°and 135°, and indicating image coarseness	RLNonUni (run length nonuniformity); GLevNonU (grey level nonuniformity); LngREmph (long run emphasis); ShrtREmp (short run emphasis); Fraction (fraction of image in runs)
Co-occurrence matrix	computed from intensities of pairs of pixels and describing the homogeneity	AngScMom (angular second moment); Contrast (contrast); Correlat (correlation); SumOfSqs (sum of squares); InvDfMom (inverse difference moment); SumAverg (sum average); SumVarnc (sum variance); SumEntrp (sum entropy); Entropy (entropy); DifVarnc (difference variance); DifEntrp (difference entropy). Features are computed for 5-pixel distance (1, 2, 3, 4, 5) and for 4 various directions (horizontal, 45 degrees, vertical, 135 degrees)
Autogressive model	coefficients of neighboring pixels reflecting the coarse-to-fine stratification	Teta1 (parameter θ1); Teta2 (parameter θ2); Teta3 (parameter θ3); Teta4 (parameter θ4); Sigma (parameter σ)
Wavelet transform	spatial frequencies at multiple scales identifying coarseness	WavEn (wavelet energy) feature is computed at 5 scales within four frequency bands: low-pass filtering in both directions (LL) assessed the lowest frequencies, low-pass filtering followed by high-pass filtering (LH) assessed horizontal edges, high-pass filtering followed by low-pass filtering (HL) assessed vertical edges; high-pass filtering in both directions (HH) assessed diagonal details.
Geometry parameters	quantifying tumor shape	GeoX (horizontal coordinate of gravity center); GeoY (vertical coordinate of gravity center); GeoF (area, number of the object pixels);

GeoSpol (diameter of the area equivalent circle);

GeoSmax (maximal diameter);

GeoAox (oriental angle);

GeoUg (specific perimeter – sum of distances between centers of neighboring contour pixels);

GeoUw (convex perimeter);

GeoFmax (maximal Feret's diameter);

GeoFmin (minimal Feret's diameter);

GeoFh (horizontal Feret's diameter);

GeoFv (vertical Feret's diameter);

GeoMmin (Martin's minimal radius);

GeoMmax (Martin's maximal radius);

GeoMaver (Martin's average radius);

GeoFt (profile area);

GeoUl (profile specific perimeter);

GeoS1 (contour-skeleton maximal thickness);

GeoS2 (contour-skeleton minimal distance);

GeoLsz (skeleton length);

GeoS (length of the circumscribing rectangle of minimal area);

GeoL (width of the circumscribing rectangle);

GeoSxL (area of circumscribing rectangle);

GeoD1 (diameter of profile inscribed circle of maximum area);

GeoD2 (diameter of circumscribing circle);

GeoFd2 (area of circumscribing circle);

GeoLmaxE (length of the circumscribing ellipsis of minimal area);

GeoLminE (width of the circumscribing ellipsis of minimal area);

GeoFE (area of circumscribing ellipsis);

GeoW1 (GeoLmaxE/GeoLminE);

GeoW2 (4 π GeoFt / GeoU1^2);

GeoW6 (1 / GeoW2);

GeoW3 (GeoU1^2 / GeoFt);

GeoW4 (GeoU1 / GeoUw);

GeoW5 (GeoF / GeoLsz);

GeoW5b (GeoLsz / GeoF);

GeoW7 (GeoD2 / GeoD1);

GeoRs (GeoU ^{2} / 4 π GeoFt);

GeoRf (GeoFh / GeoFv);

GeoRff (GeoFmax / GeoFmin);

GeoRc (circularity Rc1/Rc2);

GeoRc1 (Rc1= 2 (F/π)$^{1/2}$);

GeoRc2 (Rc2= Ug/π);

GeoRm (Malinowska ratio);

GeoRb (Blair-Bliss ratio);

GeoRd (Danielsson ratio);

GeoRh (Haralic ratio);

GeoW8 (GeoL / GeoS);

GeoW9 (GeoL GeoS / GeoF);

GeoW10 (GeoMmax / GeoMmin);

GeoSigR (standard deviation of all radii);

GeoW11 (diameter range);

GeoW12 (roundness);

GeoW13 (GeoSmax / GeoF);

GeoW14 (GeoF / GeoSmax^3);

GeoW15 (4 GeoF / π GeoFmin GeoFmax;

GeoM2x (horizontal second order moment of inertia);

GeoM2y (vertical second order moment of inertia);

GeoM2xy (second order moment of inertia);

GeoEr (average distance from gravity center)

GeoEr2 (average square distance from gravity center);

GeoEl (average distance from contour);

GeoEl2 (average square distance from contour);

GeoNc (number of contour pixels);

GeoNv (number of cavities);

GeoNl (number of profile contour pixels);

GeoNsz (number of skeleton pixels);

GeoNi (number of skeleton branches);

GeoNx (skeletal junction number);

GeoNo (number of skeletal loops)

GeoXo (gravity center to inscribed circle center horizontal distance);

GeoYo (gravity center to inscribed circle center distance);

GeoXYo (gravity center to inscribed circle center horizontal distance).

Appendix 2: default parameters of models.

No.	*Models*	*Default parameters*
1	Logistic Regression	C=1.0, class_weight=None, dual=False, fit_intercept=True, intercept_scaling=1, l1_ratio=None, max_iter=100, multi_class='auto', n_jobs=None, penalty='l2', solver='lbfgs', tol=0.0001, verbose=0, warm_start=False
2	K Neighbors Classifier	algorithm='auto', leaf_size=30, metric='minkowski', metric_params=None, n_jobs=None, n_neighbors=5, p=2, weights='uniform'
3	Naive Bayes	priors=None, var_smoothing=1e-09

4	Decision Tree Classifier	ccp_alpha=0.0, class_weight=None, criterion='gini', max_depth=None, max_features=None, max_leaf_nodes=None, min_impurity_decrease=0.0, min_impurity_split=None, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, presort='deprecated', splitter='best'
5	SVM Linear Kernel	- alpha=0.0001, average=False, class_weight=None, early_stopping=False, epsilon=0.1, eta0=0.0, fit_intercept=True, l1_ratio=0.15, learning_rate='optimal', loss='hinge', max_iter=1000, n_iter_no_change=5, n_jobs=None, penalty='l2', power_t=0.5, shuffle=True, tol=0.001, validation_fraction=0.1, verbose=0, warm_start=False
6	SVM Radial Kernel	- C=1, break_ties=False, cache_size=200, class_weight=None, coef0=0.0, decision_function_shape='ovr', degree=3, gamma='auto', kernel='rbf', max_iter=-1, probability=True, random_state=3591, shrinking=True, tol=0.001, verbose=False
7	Gaussian Process Classifier	copy_X_train=True, kernel=None, max_iter_predict=100, multi_class='one_vs_rest', n_jobs=None, n_restarts_optimizer=0, optimizer='fmin_l_bfgs_b', warm_start=False
8	MLP Classifier	activation='relu', alpha=0.0001, batch_size='auto', beta_1=0.9, beta_2=0.999, early_stopping=False, epsilon=1e-08, hidden_layer_sizes=(100,), learning_rate='constant', learning_rate_init=0.001, max_fun=15000, max_iter=500, momentum=0.9, n_iter_no_change=10,

		nesterovs_momentum=True,
		power_t=0.5, shuffle=True, solver='adam',
		tol=0.0001, validation_fraction=0.1, verbose=False,
		warm_start=False
9	Ridge Classifier	alpha=1.0, class_weight=None, copy_X=True, fit_intercept=True,
		max_iter=None, normalize=False,
		solver='auto', tol=0.001
10	Random Forest Classifier	bootstrap=True, ccp_alpha=0.0, class_weight=None,
		criterion='gini', max_depth=None, max_features='auto',
		max_leaf_nodes=None, max_samples=None,
		min_impurity_decrease=0.0, min_impurity_split=None,
		min_samples_leaf=1, min_samples_split=2,
		min_weight_fraction_leaf=0.0, n_estimators=10,
		n_jobs=None, oob_score=False,
		verbose=0, warm_start=False
11	Linear Discriminant Analysis	n_components=None, priors=None, shrinkage=None,
		solver='svd', store_covariance=False, tol=0.0001
12	Extra Trees Classifier	bootstrap=False, ccp_alpha=0.0, class_weight=None,
		criterion='gini', max_depth=None, max_features='auto',
		max_leaf_nodes=None, max_samples=None,
		min_impurity_decrease=0.0, min_impurity_split=None,
		min_samples_leaf=1, min_samples_split=2,
		min_weight_fraction_leaf=0.0, n_estimators=100,
		n_jobs=None, oob_score=False, verbose=0,
		warm_start=False

13	Extreme Gradient Boosting	base_score=0.5, booster='gbtree', colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1, gamma=0, learning_rate=0.1, max_delta_step=0, max_depth=3, min_child_weight=1, missing=None, n_estimators=100, n_jobs=-1, nthread=None, objective='binary:logistic', reg_alpha=0, reg_lambda=1, scale_pos_weight=1, seed=None, silent=None, subsample=1, verbosity=0
14	CatBoost Classifier	learning_rate(eta)=automatically, depth(max_depth)=6, l2_leaf_reg(reg_lambda)=3, n_estimators=1000, one_hot_max_size=2, loss_function='Logloss'.
15	Gradient Boosting Classifier	ccp_alpha=0.0, criterion='friedman_mse', init=None, learning_rate=0.1, loss='deviance', max_depth=3, max_features=None, max_leaf_nodes=None, min_impurity_decrease=0.0, min_impurity_split=None, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, n_estimators=100, n_iter_no_change=None, presort='deprecated', subsample=1.0, tol=0.0001, validation_fraction=0.1, verbose=0, warm_start=False
16	Ada Boost Classifier	algorithm='SAMME.R', base_estimator=None, learning_rate=1.0, n_estimators=50
17	Quadratic Discriminant Analysis	priors=None, reg_param=0.0, store_covariance=False, tol=0.0001

Appendix 3: test result of trained models based on random sampling by images

Model	Accuracy	AUC*	Recall	Precision	F1	Kappa
Logistic Regression	0.7164	0.7676	0.81	0.73	0.7679	0.4059
K Neighbors Classifier	0.82	0.8782	0.8519	0.8397	0.8457	0.6296
Naive Bayes	0.6838	0.803	0.5974	0.8065	0.6864	0.3814
Decision Tree Classifier	0.9123	0.9119	0.9211	0.9271	0.9241	0.8203
SVM - Radial Kernel	0.5606	0.5311	0.9678	0.5713	0.7185	-0.0371
Gaussian Process Classifier	0.4207	0.4689	0.0242	0.5	0.0461	-0.0077
MLP Classifier	0.7826	0.8304	0.8567	0.787	0.8204	0.5465
Random Forest Classifier	0.9356	0.9818	0.9356	0.9525	0.9439	0.8684
Linear Discriminant Analysis	0.7229	0.77	0.8148	0.7355	0.7731	0.4198
Extra Trees Classifier	0.9524	0.9412	0.9533	0.9642	0.9587	0.9026
Extreme Gradient Boosting	0.8974	0.9737	0.9114	0.9114	0.9114	0.7895
CatBoost Classifier	0.9543	0.9946	0.9549	0.9658	0.9603	0.9064
Gradient Boosting Classifier	0.9049	0.9761	0.9211	0.9152	0.9181	0.8046
Ada Boost Classifier	0.8302	0.9107	0.8631	0.8468	0.8549	0.6504
Quadratic Discriminant Analysis	0.6278	0.8427	0.4267	0.8604	0.5705	0.3026

AUC: Area Under Curve