



Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών

Παρουσίαση Διπλωματικής Εργασίας

Μελέτη manifold-based dimensionality reduction τεχνικών για ταξινόμηση αντικειμένων

Εκπόνηση:
Παπακωνσταντίνου Αντώνης
ΑΕΜ: 8977

Επίβλεψη:
Καθηγητής Α.Π.Θ.
Ιωάννης Θεοχάρης

Θεσσαλονίκη, Σεπτέμβριος 2023

Περιεχόμενα

- Μέθοδοι μείωσης διαστάσεων
- Σκοπός διπλωματικής
- Γραμμικές και μη γραμμικές μέθοδοι
- Αραιή αναπαράσταση
- Υλοποιήσεις
- Σύνολα δεδομένων
- Ταξινομητές
- Επιλογή παραμέτρων
- Αποτελέσματα πειραμάτων
- Παρατηρήσεις και συμπεράσματα
- Προβλήματα και μελλοντικές επεκτάσεις

Μέθοδοι μείωσης διαστάσεων

Κίνητρο

- Διαχείριση και ανάλυση δεδομένων υψηλής διάστασης

Οφέλη

- Αποδοτικότητα
- Οπτικοποίηση
- Αποθήκευση
- Ερμηνευσιμότητα

Μέσο

- Μετάφραση πολύπλοκων δεδομένων σε μια μορφή με βασικά χαρακτηριστικά

Κατηγορίες

- Γραμμικές μέθοδοι
- Μη γραμμικές μέθοδοι

Σκοπός διπλωματικής

- Αναδόμηση και κριτική ανάλυση δύο αλγορίθμων:
 - SPP
 - DSNPE
 - Σύγκριση με άλλους αλγορίθμους:
 - PCA
 - LPP
 - LDA
 - LLE
 - Εφαρμογή σε τρία σύνολα δεδομένων προσώπων:
 - Yale
 - Extended Yale B
 - AR
 - Ανάλυση και πειραματισμός για την αποτελεσματικότητα και αποδοτικότητα.
 - Συμβολή στην κατανόηση των τεχνικών μείωσης διάστασης.

Μέθοδοι μείωσης διαστάσεων

Γραμμικές μέθοδοι

- Μεγιστοποιεί τη διακύμανση, εντοπίζει κύριες συνιστώσες

PCA

- Αντικειμενική συνάρτηση:
- Βέλτιστο \mathbf{w} :

$$\max_{\|\mathbf{w}\|=1} \mathbf{w}^T \Sigma \mathbf{w}$$

$$\Sigma \mathbf{w} = \lambda \mathbf{w}$$

- Μεγιστοποιεί τη διάκριση των κλάσεων μέσω γραμ. συνδυασμών

LDA

- Αντικειμενική συνάρτηση:
- Βέλτιστο \mathbf{w} :

$$\max_{\mathbf{w}} \frac{\mathbf{W}^T \mathbf{S}_b \mathbf{W}}{\mathbf{W}^T \mathbf{S}_t \mathbf{W}}$$

$$\mathbf{S}_b \mathbf{w} = \lambda \mathbf{S}_t \mathbf{w}$$

- Διατηρεί τις τοπικές σχέσεις κατά την προβολή

LPP

- Αντικειμενική συνάρτηση:
- Βέλτιστο \mathbf{w} :

$$\min_{\mathbf{w}} \frac{\mathbf{w}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{w}}{\mathbf{w}^T \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{w}}$$

$$\mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{w} = \lambda \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{w}$$

Μη γραμμικές μέθοδοι

- Διατηρεί την τοπική δομή στο χαμηλότερο διάστασης χώρο

LLE

- Αντικειμενική συνάρτηση:
- Βέλτιστο \mathbf{w} :

$$\min_{\mathbf{Y}} \sum_{i=1}^N \left\| \mathbf{y}_i - \sum \mathbf{w}_{ij} \mathbf{y}_j \right\|^2$$

$$\mathbf{M} \mathbf{Y} = \lambda \mathbf{Y} \quad \text{με } \mathbf{M} = (\mathbf{I} - \mathbf{w})^T (\mathbf{I} - \mathbf{w})$$

Αραιή αναπαράσταση

Στόχος

- Να βρούμε την πιο απλή δυνατή εξήγηση ενός σήματος ως γραμμικό συνδυασμό λίγων στηλών ενός πίνακα.

Δυσκολίες

- Το πρόβλημα είναι NP-hard και υπάρχουν πολλές πιθανές λύσεις

Μέθοδος: λύση του $\min_{\mathbf{s}} \|\mathbf{s}\|_0$ subject to $\mathbf{x} = \mathbf{Xs}$

Προσέγγιση: χρήση αλγορίθμων που βρίσκουν κοντινές στη βέλτιστη λύσεις:

- Greedy αλγόριθμοι
- Bayesian στρατηγικές
- Μετατροπή του προβλήματος σε κυρτό (ℓ_1)

Υλοποίηση SPP

Είσοδος

- Δεδομένα εκπαίδευσης \mathbf{X} μεγέθους $m \times n$
- Επιθυμητή διάσταση μείωσης d
- Μια σταθερά ανοχής *epsilon*

Έξοδος

- Πίνακας προβολής μεγέθους $m \times d$

Βήματα αλγορίθμου

- 1) Προβολή το σύνολο εικόνων $\{\mathbf{x}_i\}$ στον υποχώρο PCA
- 2) Δημιουργία πίνακα βάρους \mathbf{S} :

$$\min_{\mathbf{s}_i} \|\mathbf{s}_i\|_1 \text{ subject to } \|\mathbf{x}_i - \mathbf{X}\mathbf{s}_i\| < \varepsilon, \quad \mathbf{1} = \mathbf{1}^T \mathbf{s}_i$$

- 3) Λύση του γενικευμένου προβλήματος ιδιοτιμών:

$$\mathbf{X}\mathbf{S}_\beta \mathbf{X}^T \mathbf{w} = \lambda \mathbf{X}\mathbf{X}^T \mathbf{w}, \quad \mathbf{S}_\beta = \mathbf{S} + \mathbf{S}^T - \mathbf{S}^T \mathbf{S}$$

Υλοποίηση DSNPE

Είσοδος

- Δεδομένα εκπαίδευσης \mathbf{X} μεγέθους $m \times n$
- Επιθυμητή διάσταση μείωσης d
- Μια σταθερά ανοχής *epsilon*
- Παράμετρος κανονικοποίησης γ

Έξοδος

Πίνακας προβολής μεγέθους $m \times d$

Βήματα αλγορίθμου

- 1) Προβολή το σύνολο εικόνων $\{\mathbf{x}_i\}$ στον υποχώρο PCA
- 2) Δημιουργία πίνακα βάρους \mathbf{S} :
$$\min_{\mathbf{s}_i} \|\mathbf{s}_i\|_1 \text{ subject to } \|\mathbf{x}_i - \mathbf{X}_k \mathbf{s}_i\| < \varepsilon \text{ & } \text{label}(\mathbf{x}_i) = k$$
- 3) Λύση του γενικευμένου προβλήματος ιδιοτιμών:
$$(\mathbf{X}\mathbf{S}_a\mathbf{X}^T - \gamma(\mathbf{S}_b - \mathbf{S}_w))\mathbf{w}_i = \lambda_i \mathbf{X}\mathbf{X}^T \mathbf{w}_i,$$

$$\mathbf{S}_a = \mathbf{I} - \mathbf{S} - \mathbf{S}^T + \mathbf{S}^T \mathbf{S}$$

Σύνολα δεδομένων

Yale

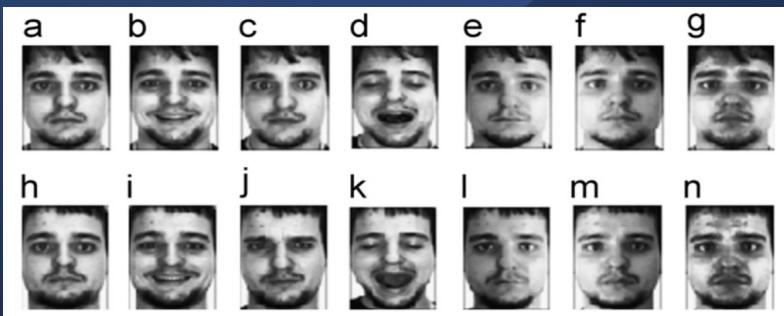
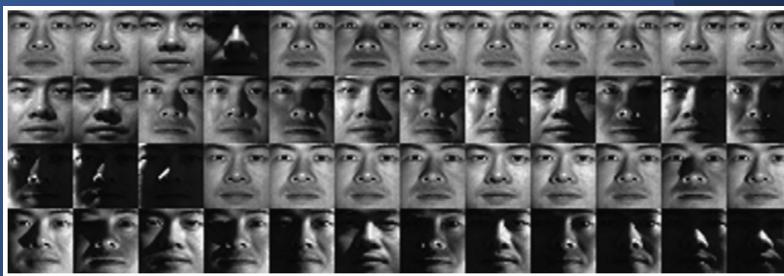
- 165 εικόνες προσώπων - 15 ατόμων
- 11 εικόνες ανά άτομο
- Διαφορετική έκφραση - κατάσταση φωτισμού
- Ασπρόμαυρες - αναδιαμορφωμένες σε 32x32 pixel

Ext. Yale B

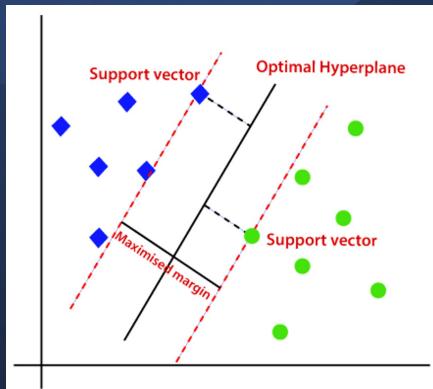
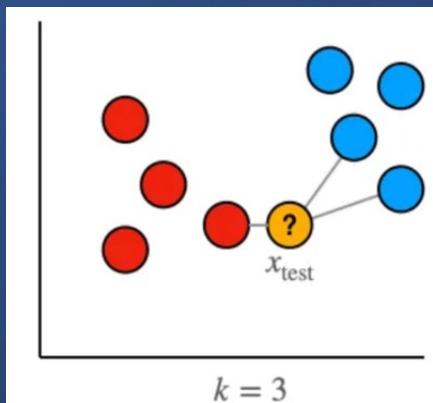
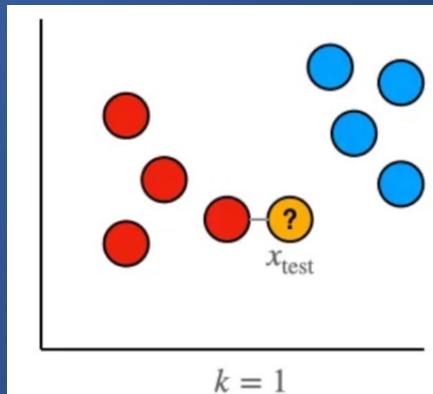
- 2414 εικόνες προσώπων - 38 ατόμων
- 64 εικόνες ανά άτομο
- Διάφορες συνθήκες φωτισμού
- Ασπρόμαυρες - αναδιαμορφωμένες σε 32x32 pixel

AR

- 1400 εικόνες προσώπων - 100 ατόμων
- 14 εικόνες ανά άτομο
- Διαφορετική έκφραση - κατάσταση φωτισμού
- Ασπρόμαυρες - αναδιαμορφωμένες σε 66x48 pixel



Ταξινομητές



Ταξινομητής 1-NN

Προσέγγιση: Κατατάσσει βάσει της κλάσης του ενός κοντινότερου γείτονα.



- Απλό και εύκολο στην υλοποίηση



- Ευαίσθητο στο θόρυβο των δεδομένων

Ταξινομητής k-NN

Προσέγγιση: Λαμβάνει υπόψη την πλειοψηφία κλάσης μεταξύ των k κοντινότερων γειτόνων.



- Δεν κάνει υποθέσεις για την κατανομή των δεδομένων



- Υπολογιστικά απαιτητικό, για μεγάλο k ή υψηλές διαστάσεις

Ταξινομητής SVM

Προσέγγιση: Βρίσκει το υπερεπίπεδο που διαιρεί καλύτερα το dataset σε κλάσεις.



- Αποτελεσματικό σε υψηλές διαστάσεις



- Απαιτείτε πολύ μνήμη για μεγάλα σύνολα δεδομένων

Επιλογή Παραμέτρων

Μέθοδοι

PCA – LDA

- Διάσταση μείωσης

LPP – LLE

- Επιλογή των γειτόνων n
- Διάσταση μείωσης

SPP

- $\epsilon_{psilon} = 0.05$
- Διάσταση μείωσης

DSNPE

- $\epsilon_{psilon} = 0.01$
- Παράμετρος κανονικοποίησης $\gamma = 1$
- Διάσταση μείωσης

Ταξινομητές

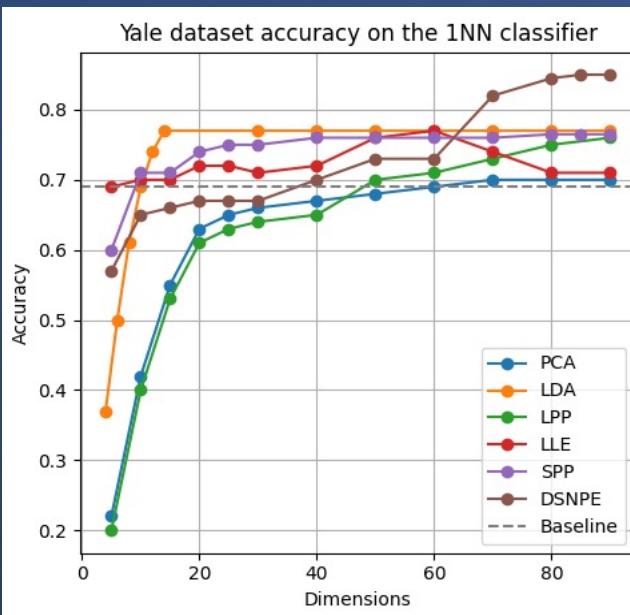
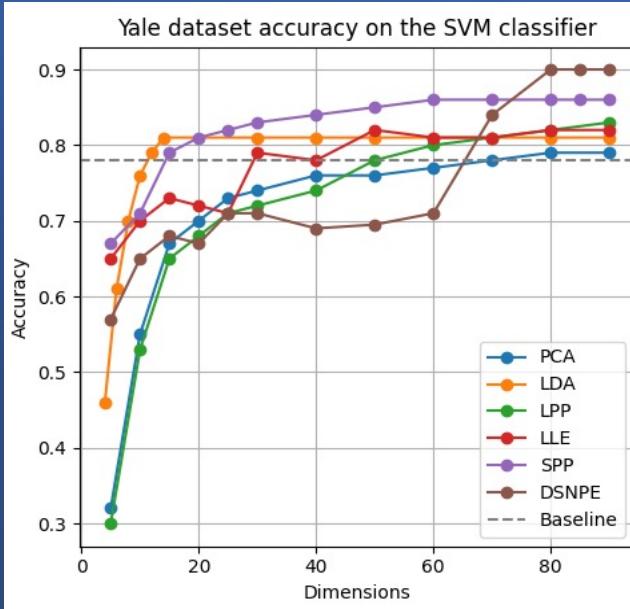
K-NN

- Cross validation με 5 folds
- Γείτονες 1 – 20

SVM

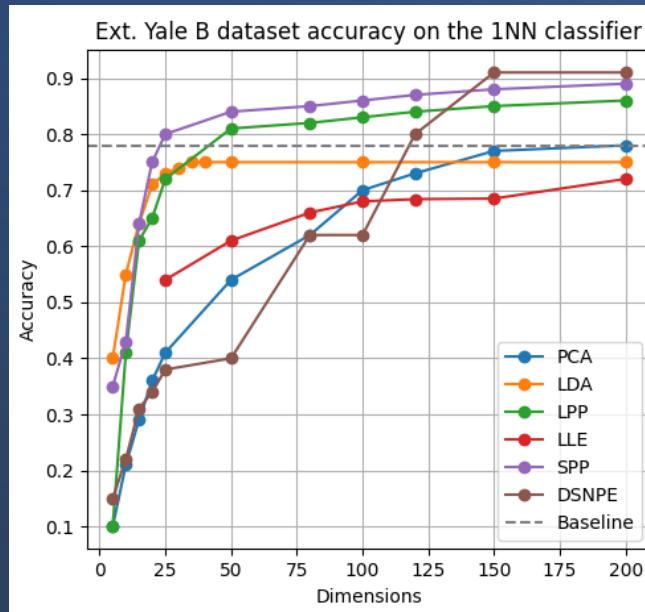
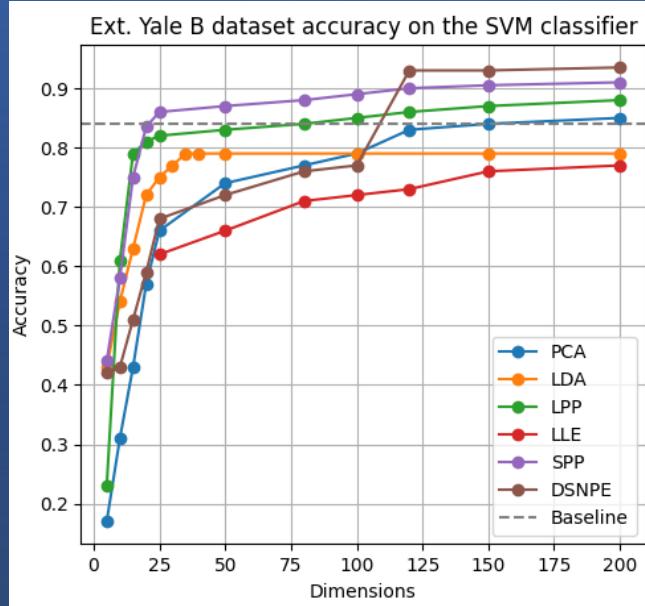
- Cross validation με 5 folds
- Kernel:
 1. ‘linear’
 2. ‘rbf’
 3. ‘poly’
 4. ‘sigmoid’

Αποτελέσματα στο Yale



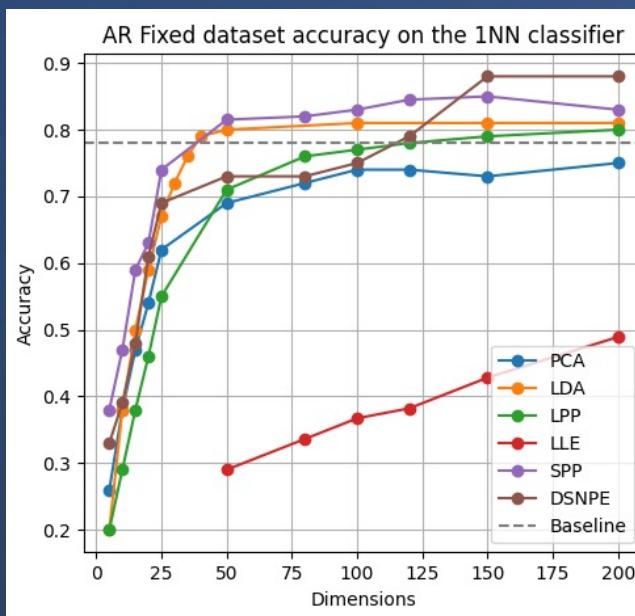
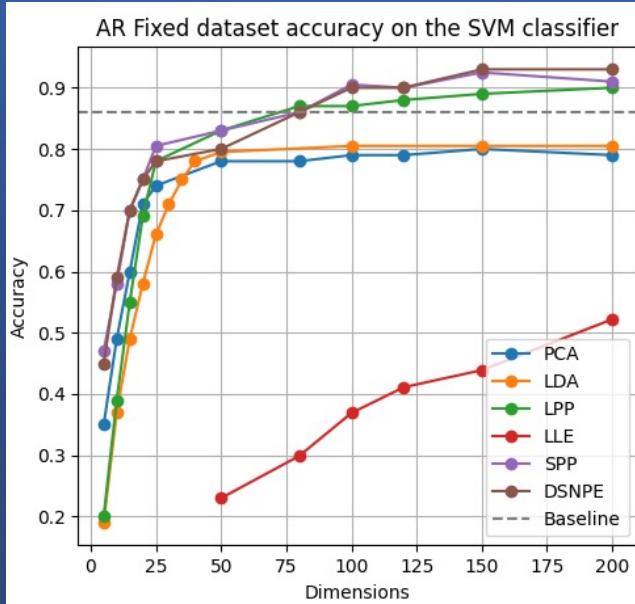
Ταξινομητής	Μέθοδος	Baseline	PCA	LDA	LPP	LLE	SPP	DSNPE
SVM	Ποσοστό % αναγνώρισης	78	79	81	83	82	86	90
	Διάσταση	1024	80	14	89	50	60	80
1-NN	Ποσοστό % αναγνώρισης	69	70	77	76	77	77	85
	Διάσταση	1024	76	14	87	60	80	82
k-NN	Ποσοστό % αναγνώρισης	69	70	77	77	78	77	85
	Διάσταση	1024	76	14	87	60	80	82

Αποτελέσματα στο Ext. Yale B



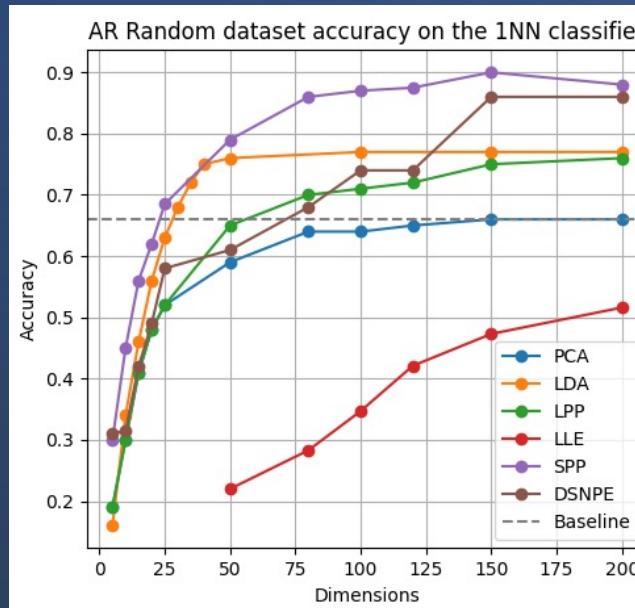
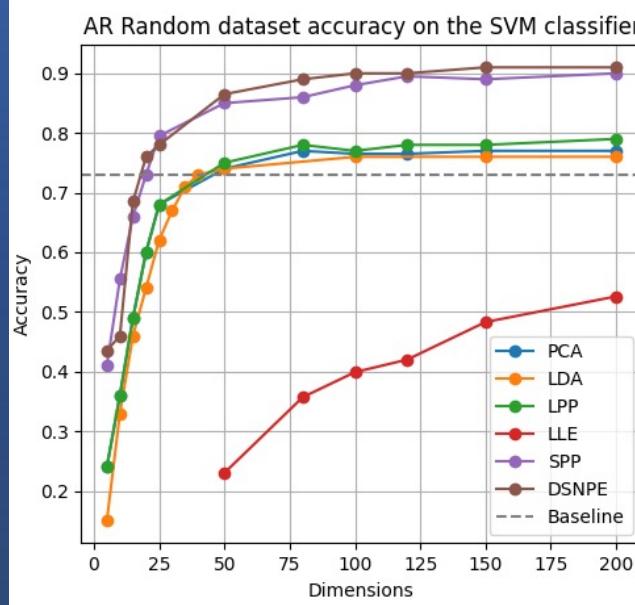
Ταξινομητής	Μέθοδος	Baseline	PCA	LDA	LPP	LLE	SPP	DSNPE
SVM	Ποσοστό % αναγνώρισης	84	85	79	88	82	91	94
	Διάσταση	1024	200	38	175	500	200	120
1-NN	Ποσοστό % αναγνώρισης	78	79	75	86	78	89	91
	Διάσταση	1024	200	36	175	500	200	150
k-NN	Ποσοστό % αναγνώρισης	78	79	76	86	78	89	91
	Διάσταση	1024	200	35	175	500	200	150

Αποτελέσματα στο AR Fixed



Ταξινομητής	Μέθοδος	Baseline	PCA	LDA	LPP	LLE	SPP	DSNPE
SVM	Ποσοστό % αναγνώρισης	86	80	80	90	80	93	93
	Διάσταση	3168	150	100	200	600	150	150
1-NN	Ποσοστό % αναγνώρισης	78	75	81	80	78	85	88
	Διάσταση	3168	200	100	200	600	150	150
k-NN	Ποσοστό % αναγνώρισης	78	75	81	80	79	85	88
	Διάσταση	3168	200	100	200	600	150	150

Αποτελέσματα στο AR Random



Ταξινομητής	Μέθοδος	Baseline	PCA	LDA	LPP	LLE	SPP	DSNPE
SVM	Ποσοστό % αναγνώρισης	73	77	76	79	81	90	91
	Διάσταση	3168	79	100	200	500	200	175
1-NN	Ποσοστό % αναγνώρισης	66	66	77	76	78	90	86
	Διάσταση	3168	150	100	200	600	150	175
k-NN	Ποσοστό % αναγνώρισης	66	66	77	76	79	90	86
	Διάσταση	3168	150	100	200	600	150	175

Παρατηρήσεις

- PCA: απλό, αλλά χαμηλή απόδοση
- LPP και LLE: βελτιωμένη απόδοση, αλλά δύσκολη επιλογή παραμέτρων
- SPP και DSNPE: υψηλή απόδοση με διάκριση, χωρίς προσαρμογή παραμέτρων
- DSNPE και SVM: καλύτερος συνδυασμός
- DSNPE: υψηλή διάσταση για υψηλή απόδοση
- SPP: γρήγορο και συμπαγές

Συμπεράσματα

- Επιτυχής ανακατασκευή δύο αλγορίθμων: SPP και DSNPE.

- Και οι δύο αλγόριθμοι δίνουν προτεραιότητα στη διατήρηση αραιών βαρών ανακατασκευής.

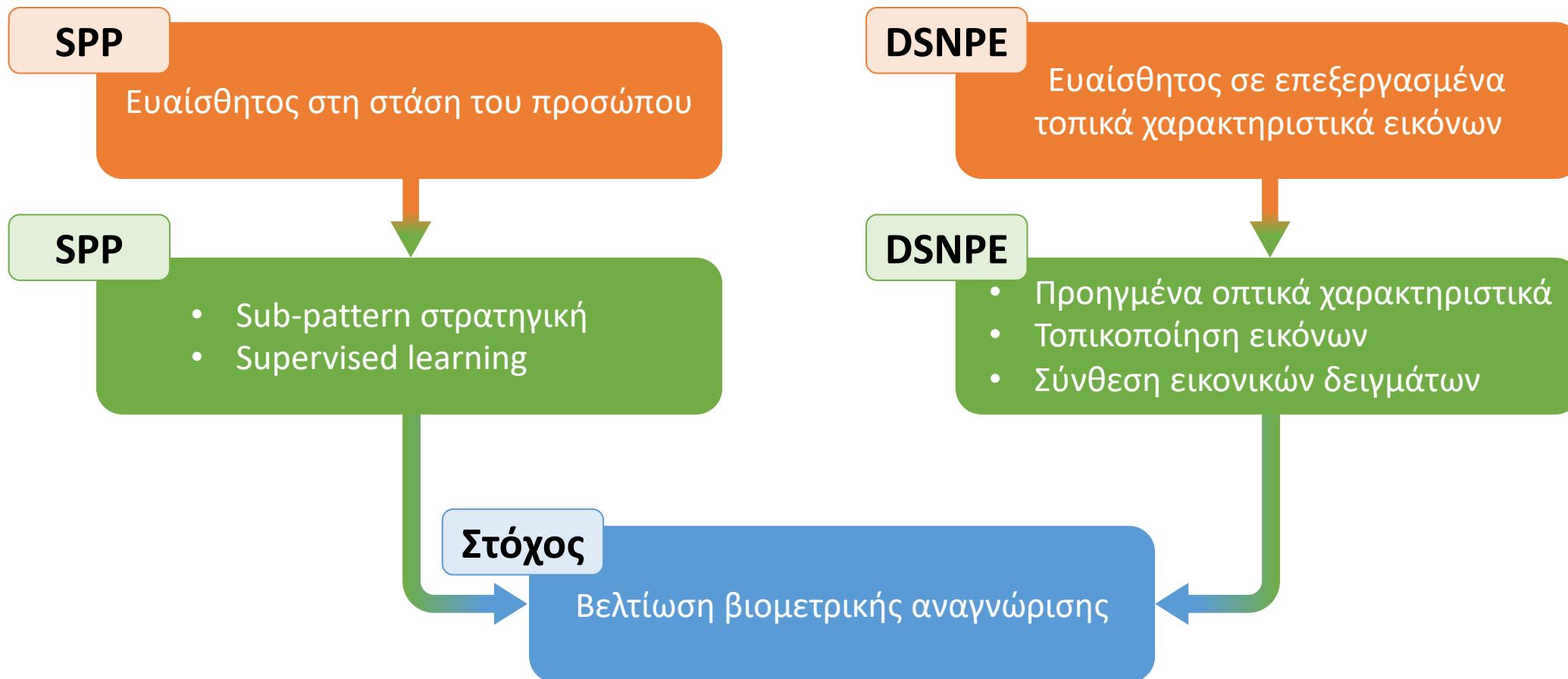
- Ο SPP υπερτερεί των PCA, LDA, LPP και LLE.

- Το SPP μπορεί να αντιμετωπίσει την ασάφεια στον προσδιορισμό της “τοπικότητας”, του LPP.

- Ο DSNPE συνδυάζει ανθεκτικότητα αραιής αναπαράστασης και διακριτικότητα MMC.

- Η μέθοδος DSNPE υπερτερεί σε όλα τα σύνολα δεδομένων.

Προβλήματα και μελλοντικές επεκτάσεις



Βιβλιογραφία

- S. Roweis, L. Saul, Nonlinear dimensionality reduction by locally linear embedding, *Science* 290 (5500) (2000) 2323–2326.
- X. He, P. Niyogi, Locality preserving projections, in: *Proceedings of Conference on Advances in Neural Information Processing Systems (NIPS)*, 2003.
- J. Wright, A. Yang, S. Sastry, Y. Ma, Robust face recognition via sparse representation, *IEEE Trans. Pattern Anal. Mach. Intell.* 31 (2) (2009) 210–227.
- M. Turk, A. Pentland, Eigenfaces for recognition, *J. Cognitive Neurosci.* 3 (1)(1991) 71–86.
- M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, *IEEE Trans. Image Process.* 15 (12) (2006) 3736–3745.
- H. Li, T. Jiang, K. Zhang, Efficient and robust feature extraction by maximum margin criterion, *IEEE Trans. Neural Networks* 17 (1) (2006) 157–165.
- B. Scholkopf, A. Smola, K. Muller, Kernel principal component analysis, in: B. Scholkopf, C. Burges, A. Smola (Eds.), *Advances in Kernel Methods—Support Vector Learning*, MIT Press, Cambridge, MA, 1999, pp. 327–352.
- K. Huang, S. Aviyente, Sparse representation for signal classification, in: *Advances in Neural Information Processing Systems (NIPS)*, 2006.
- X. He, S. Yan, Y. Hu, P. Niyogi, H. Zhang, Face recognition using Laplacianfaces, *IEEE. Trans. Pattern Anal. Mach. Intell.* 27 (3) (2005).

Σας ευχαριστώ πολύ.

Ερωτήσεις;