

Νευρωνικά Δίκτυα - Βαθιά Μάθηση

Παρουσίαση Υποχρεωτικών Εργασιών

Κωνσταντής Αθανάσιος 10537

Aristotle University of Thessaloniki
Department of Electrical and Computer Engineering
Electronics Division

Ιανούαριος 2025

Για τα παρακάτω πειράματα χρησιμοποιήθηκε η βιβλιοθήκη
Cifar-10

KNN-Nearest Centroid

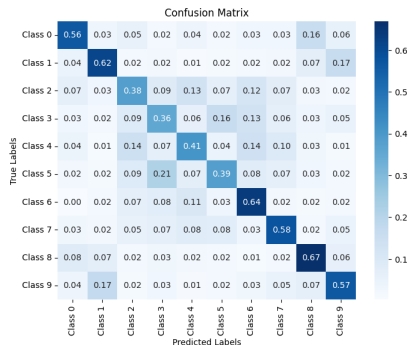
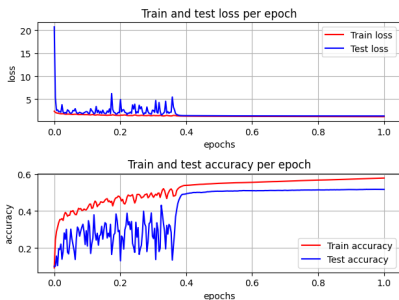
Η ενδιαμέση εργασία ζητούσε την υλοποίηση των αλγορίθμων KNN και Nearest Centroid.

Υλοποιήσαμε δικές μας συναρτήσεις και για τους δυο αυτούς αλγορίθμους οι οποίες ήταν παραλληλοποιημένες. Συνεπώς, οι υπολογισμοί έπαιρναν λίγα μόνο δευτερόλεπτα.

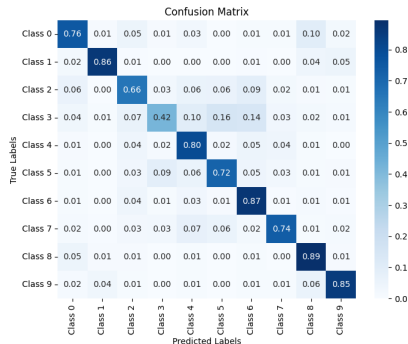
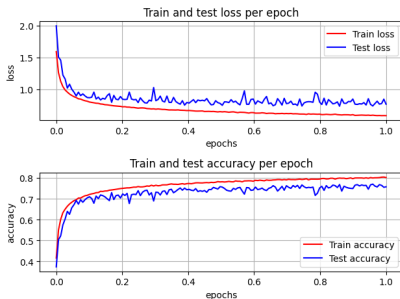
	PIXL	HIST	HOG	EDGE	IMAGE
1-NN	35.39	26.72	49.02	16.70	11.55
3-NN	33.03	36.43	49.86	15.30	10.42
NC	27.74	26.72	49.02	16.70	11.55

MLP

Για την εργασία 1 δοκιμάστηκαν διάφορα MLPs με πολλαπλά Linear Layers, Batch Normalization και χρησιμοποιώντας για activation function τη ReLU. Έγινε training με batches και χωρίς. Τελικά καταλήξαμε σε μέγιστο ποσοστό επιτυχίας 52% με raw data και 62% με HOG feature extraction. Οι χρόνοι εκπαίδευσης εξαρτώνταν από την περιπλοκότητα του μοντέλου και ήταν μόλις μερικά λεπτά.



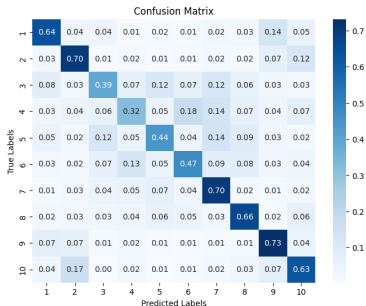
Δοκιμάστηκε επίσης και η δημιουργία CNN το οποίο θεωρητικά λειτουργεί πολύ καλύτερα στην κατηγοριοποίηση εικόνων. Επίσης, εφαρμόστηκε cross validation στο καλύτερο μοντέλο CNN για να βελτιστοποιηθεί ακόμη περισσότερο. Τα CNNs παίρνουν πολλή υπολογιστική ισχύ και χρειάζονται μεγάλη VRAM στην GPU μας, οπότε κάναμε το training με batches. Το μεγαλύτερο ποσοστό που καταφέραμε ήταν το 75% μετά από πολλές ώρες εκτέλεσης του αλγορίθμου.



SVMs

- Για τα SVMs χρησιμοποιήθηκε η βιβλιοθήκη scikit-learn.
- Κάναμε διάφορα parameter searches, και επειδή η scikit-learn τρέχει στην cpu, αυτά έπαιρναν πολλαπλές ώρες (επειδή υλοποιούσαν και cross validation).
- Γενικά τα SVMs για high dimensionality στο input space παίρνουν πολύ χρόνο και δεν είναι πολύ αποδοτικά.
- Εφαρμόσαμε PCA στα δεδομένα κρατώντας το 90% της πληροφορίας.

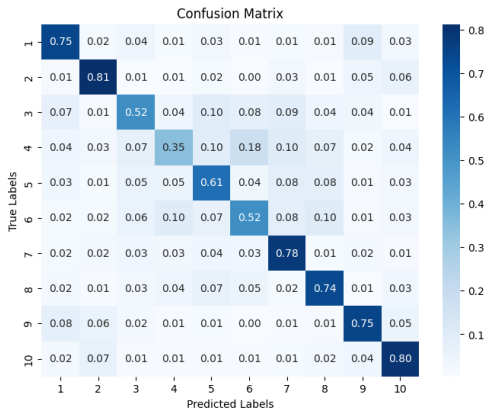
Το καλύτερο accuracy που βρέθηκε ήταν 56.9%.



Το καλύτερο accuracy που βρέθηκε ήταν 56.9% με rbf kernel.

Αφού τα HOG features δούλεψαν τόσο καλά στο παρελθόν, δοκιμάσαμε να τα χρησιμοποιήσουμε και εδώ.

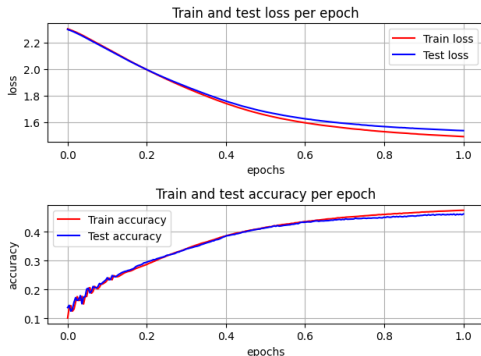
Τρέξαμε επίσης πολλαπλά grid searches με διάφορα linear, rbf και polynomial kernels διαφορετικών βαθμών.



Καταλήξαμε και πάλι σε rbf kernel με 66.27% accuracy.

Τα SVMs τελικά είναι κάπως αδύναμα στη cifar-10 αλλά υπερτερούν σε σχέση με ένα single layer MLP με Hinge Loss και με τον 1 κ 3-NN.

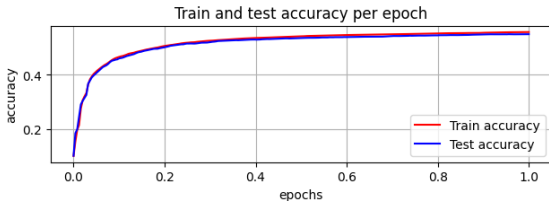
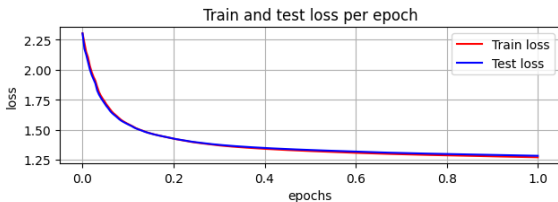
- Υλοποιήσαμε με την PyTorch ένα RBF NN με μεταβλητό αριθμό kernels.
- Του δώσαμε τα raw data από την cifar-10.
- Δοκιμάστηκε εκπαίδευση με raw data αλλά και με PCA κρατώντας 90% της πληροφορίας.
- Δοκιμάστηκε η χρήση K-Means και Centroids για την αρχικοποίηση των kernel functions.



Η βέλτιστη λύση ήταν αυτή με 50 kernels, χρησιμοποιώντας raw data χωρίς PCA. Το accuracy της λύσης έφτασε το 46.10%.

RBF & HOG

Με τη χρήση HOG Features, με 50 kernels, φτάνουμε το accuracy του 55%. Αυτό είναι λογικό μιας και η γεωμετρική απόσταση στα HOG Features, από τον K-NN καταλάβαμε ότι μπορεί να χρησιμοποιηθεί με μεγαλύτερη ακρίβεια για την κατηγοριοποίηση.



- Τα CNNs είναι τα βέλτιστα μοντέλα για κατηγοριοποίηση datasets εικόνων όπως το cifar-10. Απαιτούν όμως πολύ χρόνο και πολλή υπολογιστική ισχύ.
- Σε κάθε άλλη περίπτωση, η χρήση των HOG features οδηγεί σε καλύτερα αποτελέσματα από τη χρήση raw data, και η χρονική καθυστέρηση για τον υπολογισμό τους τελικά αξίζει συνολικά.