

Object Classification For Autonomous Driving

Machine Learning

Ηρακλής Σπύρου
Αλέξανδρος Νταγιάντας

Εισαγωγή-Στόχοι

- ◆ **Σημαντικότητα της Αυτόνομης Οδήγησης:**

Η ακριβής ταξινόμηση αντικειμένων είναι κρίσιμη για την αυτόνομη οδήγηση, καθώς μέσω της αντίληψης και της αλληλεπίδρασης τα οχήματα λαμβάνουν αποφάσεις που βελτιώνουν την ασφάλεια, αυξάνουν την αποδοτικότητα και προσφέρουν μεγαλύτερη άνεση στις μεταφορές.

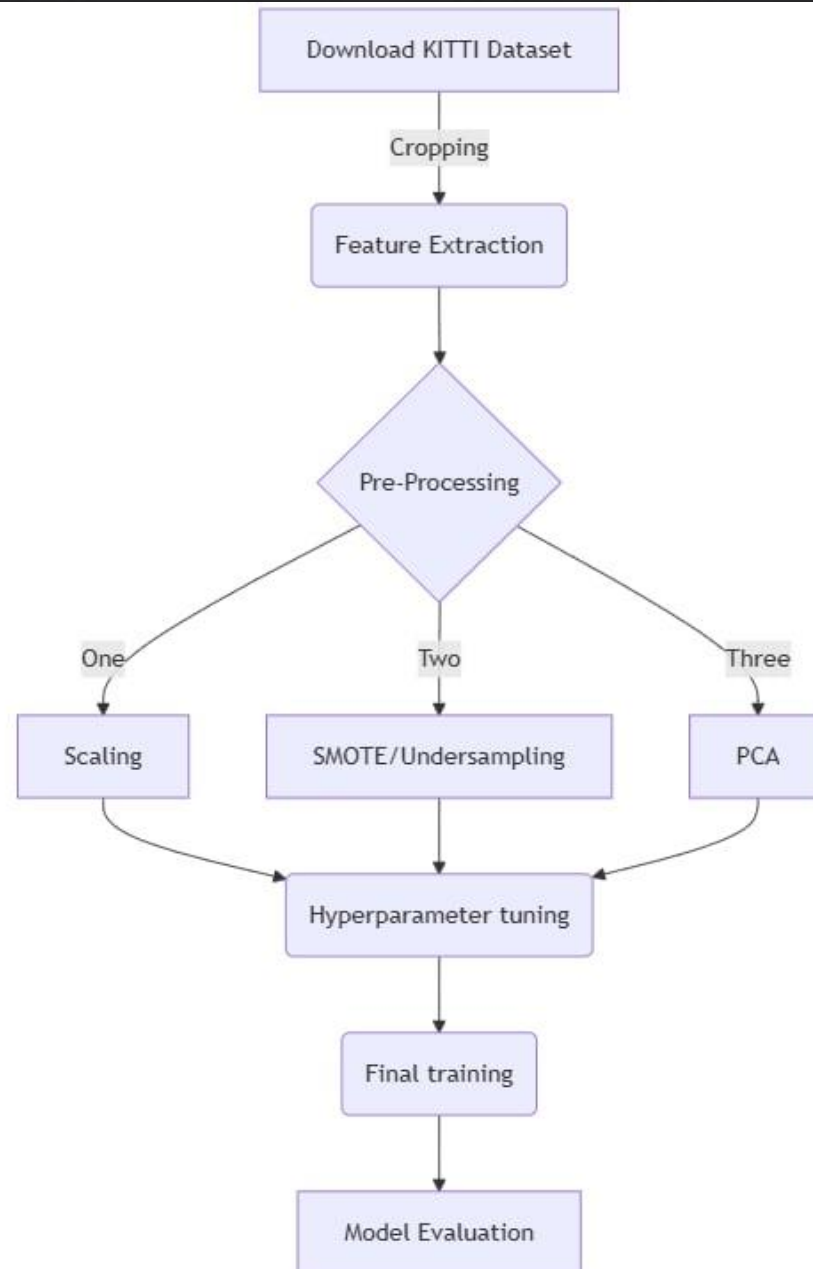
- ◆ **Χρήση του KITTI Dataset:**

Το report βασίζεται στο KITTI dataset, το οποίο παρέχει έγχρωμες εικόνες με αντικείμενα μαζί με τα αντίστοιχα labels, προσφέροντας μια αξιόπιστη βάση δεδομένων για την ανάλυση.

- ◆ **Στόχος της εργασίας:**

Η σύγκριση τεσσάρων μοντέλων ταξινόμησης (SVM, KNN, GNB, RF) ως προς την ικανότητα ταξινόμησης των κλάσεων Car, Pedestrian, Cyclist, Tram.

Workflow



KITTI Dataset

Train+Test Dataset:

- ❖ *data_object_image_2.zip*:
Περιέχει **έγχρωμες** εικόνες για το **2D-object dataset**, συνολικού μεγέθους **12GB**.
- ❖ *data_object_label.zip*:
Περιέχει τα labels για το training set, συνολικού μεγέθους **5MB**.
- ❖ Συνολικά **7481 εικόνες** (πρώτο sub-folder)

Παράμετρος	Περιγραφή
Object Type	Η κλάση του αντικειμένου.
Truncation	Το επίπεδο περικοπής του αντικειμένου.
Occlusion Alpha	Το επίπεδο απόκρυψης του αντικειμένου από άλλα αντικείμενα.
Bounding Box	Ο παρατηρούμενος προσανατολισμός του αντικειμένου.
3D Dimensions	Οι συντεταγμένες του ορθογώνιου περιγράμματος (xmin, ymin, xmax, ymax).
3D Location	Οι διαστάσεις του αντικειμένου στο 3D χώρο (height, width, length).
Rotation Y	Η θέση του αντικειμένου στο 3D χώρο (x, y, z).
	Ο προσανατολισμός του αντικειμένου γύρω από τον κάθετο άξονα.



Cropping

◈ Σκοπός του Cropping:

- Απομόνωση αντικειμένων ενδιαφέροντος
- Μείωση background noise

◈ Διαδικασία:

- Ανάγνωση εικόνων & labels (bounding boxes)
- Εξαγωγή Bounding Boxes (xmin, ymin, xmax, ymax)
- Εφαρμογή Cropping με OpenCV
- Αποθήκευση cropped εικόνων σε ξεχωριστό φάκελο

◈ Κριτήρια Φιλτραρίσματος:

- Αγνόηση πολύ κομμένων ή ελάχιστα ορατών αντικειμένων
- Συμπερίληψη μόνο των κλάσεων: Car, Pedestrian, Cyclist, Tram

◈ Συνολικά: μετά το cropping: 24614 εικόνες



Εξαγωγή Χαρακτηριστικών

◈ Διαδικασία:

- Μετατροπή σε Grayscale
- Resize (128×64 pixels)
- Εφαρμογή HOG (από OpenCV) για υπολογισμό της κλίσης και του προσανατολισμού των άκρων
- Αποθήκευση σε αρχείο .csv

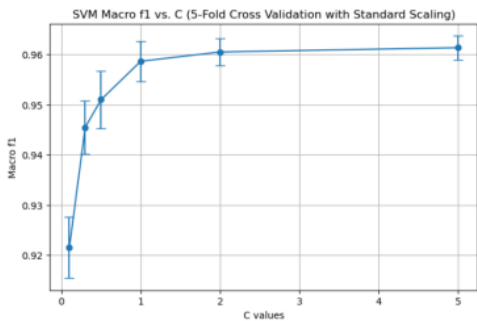
◈ Τελικά: Feature vector (3780 χαρακτηριστικά)×24614 instances

Διαδικασία Εκπαίδευσης

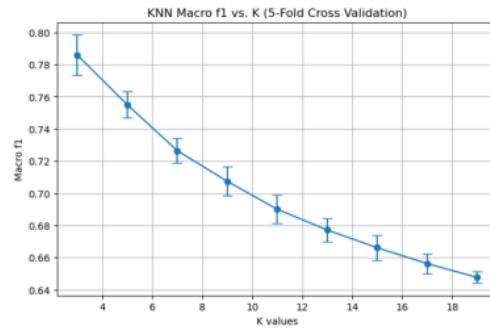
- ♦ **Διαχωρισμός Δεδομένων:** training+validation και testing.
- ♦ **Scaling:** Standard Scaler (mean=0, std=1).
- ♦ **Εξισορρόπηση Δεδομένων:**
Χρήση τεχνικών **SMOTE** για δημιουργία συνθετικών δεδομένων και **Undersampling** για την αφαίρεση δεδομένων με ομοιόμορφο τρόπο.
- ♦ **Μείωση Διαστατικότητας (PCA):** (n_components=300).
- ♦ **Hyperparameter tuning:**
Εκπαίδευση των μοντέλων με 5-fold cross validation και επιλογή βέλτιστων υπερπαραμέτρων.
- ♦ **Επανάληψη της διαδικασίας:** αλλά χωρίς dataset balancing για σύγκριση αποτελεσμάτων.
- ♦ **Τελική εκπαίδευση:** σε όλο το train+val dataset (χρησιμοποιώντας τη βέλτιστη υπερπαραμέτρο) και αξιολόγηση στο test dataset.

Ταξινομητής	Default Τιμές	Πλέγμα Τιμών
SVM	C=1.0	[0.1, 0.3, 0.5, 1.0, 2.0, 5.0]
KNN	n_neighbors=5	[3, 5, 7, 9, 11, 13, 15, 17, 19]
GNB	var_smoothing=1e-9	logspace(-5, -14, num=10)
Random Forest	n_estimators=100	[100, 150, 200, 250, 300]

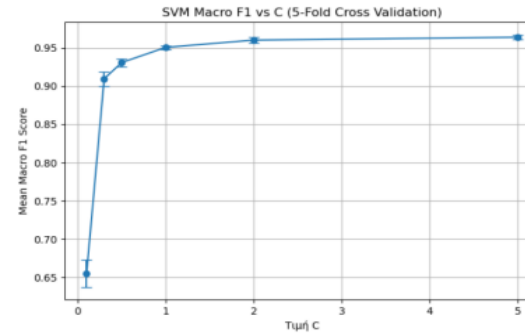
Αποτελέσματα Hyperparameter Tuning



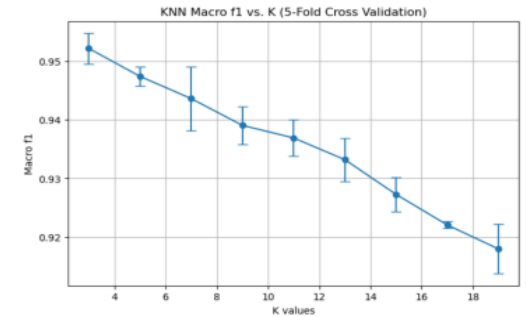
(a) SVM (Balanced)



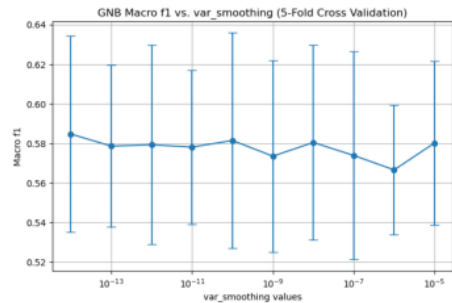
(b) KNN (Balanced)



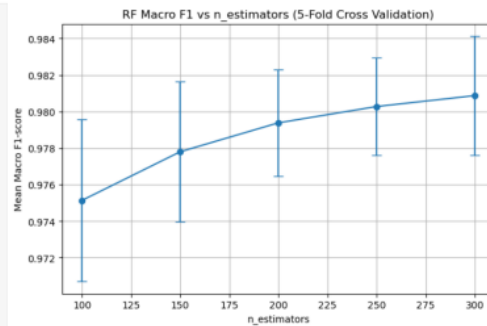
(a) SVM (Unbalanced)



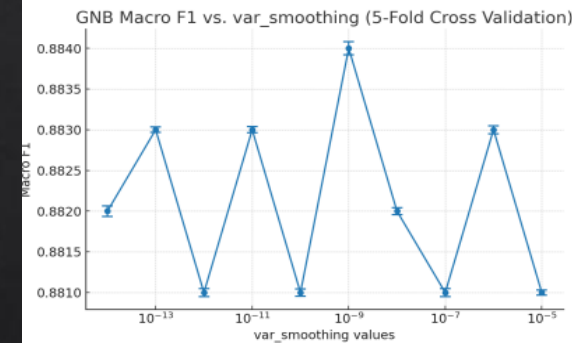
(b) KNN (Unbalanced)



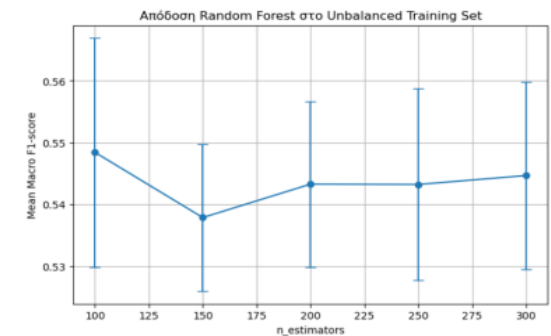
(c) GNB (Balanced)



(d) Random Forests (Balanced)



(c) GNB (Unbalanced)



(d) Random Forests (Unbalanced)

Μοντέλο	Τελική Τιμή Υπερπαραμέτρου	Balanced
SVM	$C = 2.0$	NAI
KNN	$n_neighbors = 3$	OXI
GNB	$var_smoothing = 10^{-11}$	OXI
RF	$n_estimators = 300$	NAI

Τελικά αποτελέσματα (Classification Report & Confusion Matrices)

Table 6: Classification Report – SVM

	Prec.	Rec.	F1	Sup.
Car	1.00	0.99	1.00	3890
Cyclist	0.94	0.91	0.92	241
Pedestrian	0.95	0.99	0.97	734
Tram	1.00	1.00	1.00	58
Accuracy	-	-	0.99	4923
Macro avg	0.97	0.97	0.97	4923
Weighted avg	0.99	0.99	0.99	4923

Table 7: Classification Report – KNN

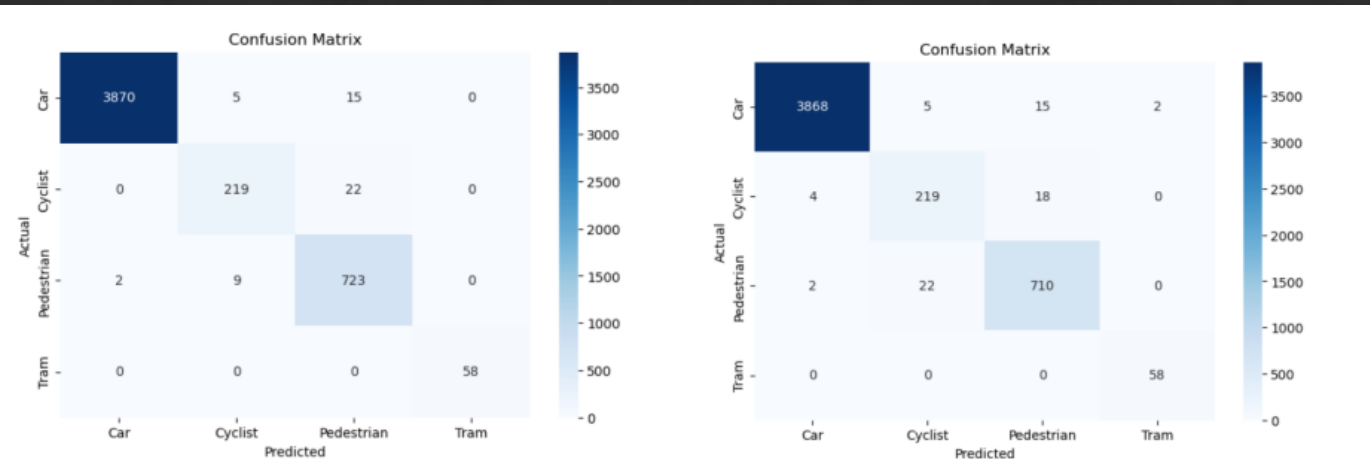
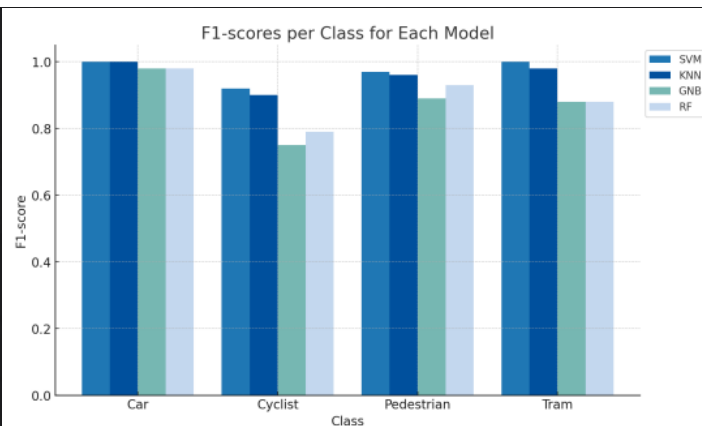
	Prec.	Rec.	F1	Sup.
Car	1.00	0.99	1.00	3890
Cyclist	0.89	0.91	0.90	241
Pedestrian	0.96	0.97	0.96	734
Tram	0.97	1.00	0.98	58
Accuracy	-	-	0.99	4923
Macro avg	0.95	0.97	0.96	4923
Weighted avg	0.99	0.99	0.99	4923

Table 8: Classification Report – GNB

	Prec.	Rec.	F1	Sup.
Car	0.98	0.99	0.98	3890
Cyclist	0.80	0.70	0.75	241
Pedestrian	0.91	0.87	0.89	734
Tram	1.00	0.79	0.88	58
Accuracy	-	-	0.96	4923
Macro avg	0.92	0.84	0.88	4923
Weighted avg	0.96	0.96	0.96	4923

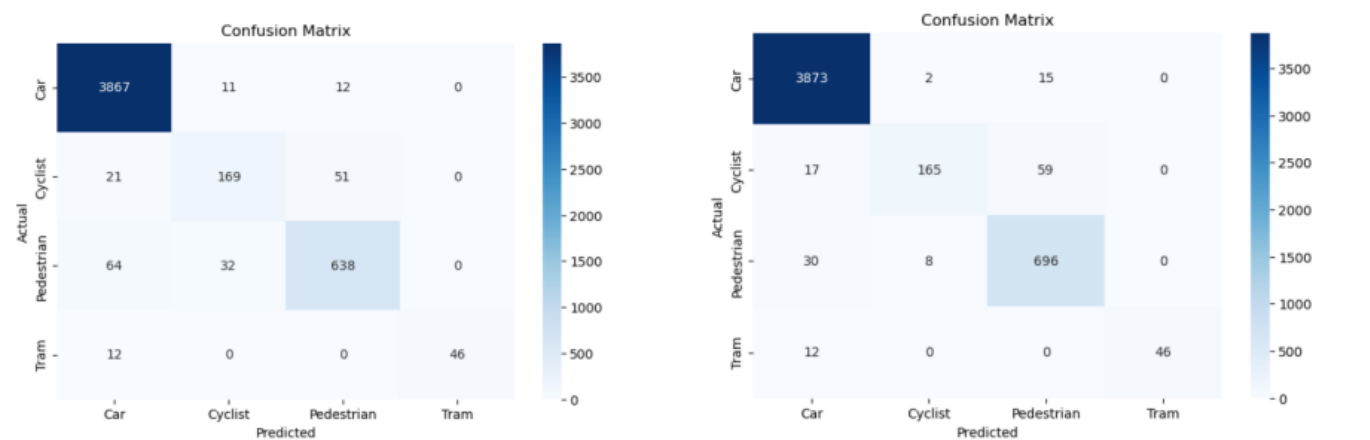
Table 9: Classification Report – RF

	Prec.	Rec.	F1	Sup.
Car	0.98	1.00	0.99	3890
Cyclist	0.94	0.68	0.79	241
Pedestrian	0.90	0.95	0.93	734
Tram	1.00	0.79	0.88	58
Accuracy	-	-	0.97	4923
Macro avg	0.96	0.86	0.90	4923
Weighted avg	0.97	0.97	0.97	4923



(a) Confusion Matrix – SVM (Test Set)

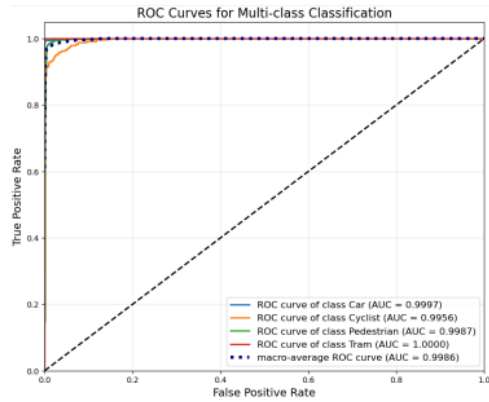
(b) Confusion Matrix – KNN (Test Set)



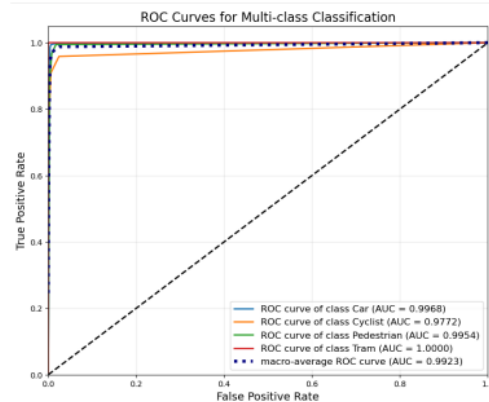
(c) Confusion Matrix – GNB (Test Set)

(d) Confusion Matrix – Random Forest (Test Set)

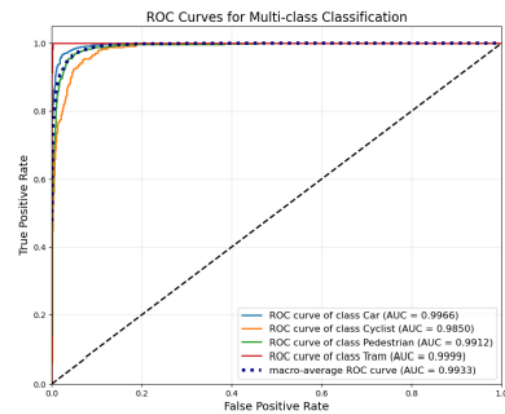
Τελικά αποτελέσματα (ROC & AUC)



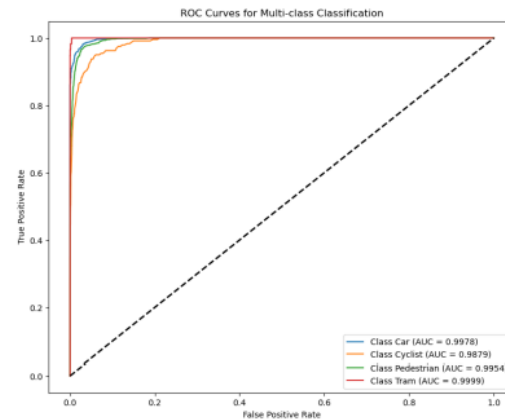
(a) Καμπύλη ROC για το SVM



(b) Καμπύλη ROC για το KNN



(c) Καμπύλη ROC για το GNB



(d) Καμπύλη ROC για το RF

Μοντέλο	Car	Cyclist	Pedestrian	Tram
SVM	0.9997	0.9956	0.9987	1.0000
KNN	0.9968	0.9772	0.9954	1.0000
GNB	0.9966	0.9850	0.9912	0.9999
RF	0.9978	0.9879	0.9954	0.9999

Table 10: AUC Scores για κάθε κλάση ανά μοντέλο.

Τελικά Συμπεράσματα

SVM:

- Εμφάνισε υψηλή ακρίβεια και σταθερότητα είτε στο balanced είτε στο unbalanced dataset.
- Είχε το μεγαλύτερο recall σε όλες τις κλάσεις και τα πήγε εξαιρετικά καλά στην κλάση cyclist.

KNN

- Είχε αρκετά καλύτερη απόδοση στο unbalanced dataset
- Παρουσιάζει ανταγωνιστική απόδοση, με ελαφρώς μειωμένη ακρίβεια συγκριτικά με το SVM.

GNB

- Αρκετά καλύτερη απόδοση στο unbalanced
- Εμφανίζει χαμηλότερη ακρίβεια σε σχέση με τα υπόλοιπα μοντέλα, κυρίως στις κατηγορίες Cyclist και Tram

RF

- Αρκετά καλύτερη απόδοση στο balanced dataset
- Αντιμετωπίζει προκλήσεις στην ταξινόμηση Cyclist και Tram, πιθανότατα λόγω της πολυπλοκότητας του dataset.

Ερωτήσεις;



ANY
QUESTIONS
?