Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Σημάτων, Ελέγχου και Ρομποτικής
Εργαστήριο Όρασης Υπολογιστών, Επικοινωνίας Λόγου και Επεξεργασίας Σημάτων

Αναγνώριση και εντοπισμός ανθρώπινης δραστηριότητας σε βίντεο

Διπλωματική εργασία

του

Ευστάθιου Ε. Γαλανάκη

Επιβλέπων: Πέτρος Μαραγκός
Καθηγητής Ε.Μ.Π.

Αθήνα, Νοέμβριος 2019

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Σημάτων, Ελέγχου και Ρομποτικής
Εργαστήριο Όρασης Υπολογιστών, Επικοινωνίας Λόγου και Επεξεργασίας Σημάτων

Αναγνώριση και εντοπισμός ανθρώπινης δραστηριότητας σε βίντεο

Διπλωματική εργασία

του

Ευστάθιου Ε. Γαλανάκη

Επιβλέπων: Πέτρος Μαραγκός
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την - 2019.

(Υπογραφή)	*(Υπογραφή)*	*(Υπογραφή)*
.....	.....	.....
Πέτρος Μαραγκός	-	-
Καθηγητής	-	-
Ε.Μ.Π	-	-

Αθήνα, Νοέμβριος 2019

(Υπογραφή)

.....
Ευστάθιος Ε. Γαλανάκης
Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Ευχαριστίες

-

Περίληψη

Σκοπός αυτής της διπλωματικής εργασίας είναι ο σχεδιασμός ενός δικτύου αναγνώρισης και εντοπισμού των ανθρώπινων ενεργειών σε βίντεο. Το δίκτυό μας στοχεύει να προσδιορίσει χωροχρονικά μια αναγνωρισμένη ενέργεια μέσα σε ένα βίντεο παράγωντας μια ακολουθία δισδιάστατων κουτιών, ένα για κάθε εικόνα του βίντεο, που περικλείει το άτομο που εκτελεί την αναγνωρισμένη ενέργεια.

Η ανίχνευση και η αναγνώριση των ενεργειών σε βίντεο είναι μια από τις μεγαλύτερες προκλήσεις στο πεδίο της Όρασης Υπολογιστών. Οι πιο πρόσφατες προσεγγίσεις περιλαμβάνουν ένα δίκτυο ανίχνευσης αντικειμένων το οποίο προτείνει δυσδίαστα κουτάκια ανά εικόνα, έναν αλγόριθμο σύνδεσης για τη δημιουργία υποψήφιων action tubes και έναν ταξινομητή για την ταξινόμησή τους. Πάνω σάυτό, οι περισσότερες από αυτές τις προσεγγίσεις εξαγάγουν τις χρονικές πληροφορίες από ένα δίκτυο το οποίο εκτιμά οπτική ροή σε επίπεδο πλαισίου. Η εισαγωγή των τρισδιάστατων συνελικτικών δικτύων μας έχει βοήθησει να μπορούμε να υπολογίσουμε τις χωροχρονικές πληροφορίες από τα βίντεο και ταυτόχρονα να εξάγουμε χωροχρονικά χαρακτηριστικά. Η προσέγγισή μας προσπαθεί να συνδυάσει τα οφέλη του να χρησιμοποιείς δίκτυα ανίχνευσης αντικειμένων και τρισδιάστατες συνελίξεις. Σχεδιάζουμε ένα δίκτυο του οποίου η δομή βασίζεται στα κλασσικά δίκτυα εντοπισμού δράσης και το ονομάζουμε ActionNet. Το πρώτο στοιχείο είναι ένα τρισδιάστατο ResNet34 το οποίο χρησιμοποιείται για τη χωροχρονική εξαγωγή χαρακτηριστικών. Επίσης, σχεδιάζουμε ένα δίκτυο για προτείνει υποψίφιες ακολουθίες από δισδιάστατα κουτιά με βάση τα χωροχρονικά χαρακτηριστικά, που το ονομάζουμε Tube Proposal Network. Αυτό το δίκτυο είναι μια επέκταση του Region Proposal Network παίρνοντας ως είσοδο τα εξαγώμενα χαρακτηριστικά και δίνοντας ως εξόδων k προτεινόμενες ακολουθίες από δισδιάστατα κουτιά. Εξετάζουμε 2 προσεγγίσεις για τον καθορισμό των τρισδιάστατων προκαθορισμένων κουτιών(anchors), τα οποία χρησιμοποιεί το TPN. Επιπλέον, σχεδιάζουμε έναν αλγόριθμο σύνδεσης για τη σύνδεση των προτεινόμενων σωλήνων δράσης. Τέλος, διερευνούμε αρκετές τεχνικές ταξινόμησης, συμπεριλαμβανομένου ενός ταξινομητή SVM, ενός Linear, ενός RNN και ενός MLP για τα σύνολα δεδομένων JHMDB και UCF101.

Λέξεις κλειδία

-

Abstract

The purpose of this diploma thesis is the design of a network for recognizing and localising human actions in videos. Our network aims to spatio-temporally localize a recognized action within a video producing a sequence of 2D boxes, one per frame, which includes the actor performing the recognized action.

Detecting and Recognizing actions in videos is one of the biggest challenges in the field of Computer Vision. Most recent approaches includes an object detection network which proposes bounding boxes per frame, a linking method for creating candidate action tubes and a classifier for classifying these. On top of that, most of these approaches extract temporal information from a network which estimates optical flow in frame level. The introduction of 3D Convolutional Networks has helped us estimating spatio-temporal information from videos and simultaneously extract spatio-temporal features. Our approach tries to combine the benefits from using object detection networks and 3D Convolution.

We design a network whose structure is based on standard action localization networks and we name it ActionNet. Its first element is a 3D ResNet34 which is used for spatio-temporal feature extraction. Also, we design a network for proposing action tubes based on spatio-temporal features, called Tube Proposal Network. This network is an expansion of Region Proposal Network and it gets as input the extracted features and outputs k-proposed action tubes. We explore 2 approaches for defining 3D anchors, which TPN uses. On top of that, we design a linking algorithm for connecting proposed action tubes. Finally, we explore several classification techniques including a SVM classifier and a MLP for datasets JHMDB and UCF101.

Keywords

Action Localization, Action Recognition, Action Tubes

Περιεχόμενα

Κατάλογος Πινάκων

Κατάλογος Σχημάτων

Κεφάλαιο 1

Εισαγωγή

Στις μέρες μας, η τεράστια αύξηση της υπολογιστικής ισχύος των Η/Υ μας βοηθά να αντιμετωπίσουμε πολλές δύσκολες καταστάσεις που εμφανίζονται στην καθημερινότητά μας. Πολλοί τομείς της επιστήμης κατάφεραν να αντιμετωπίσουν σημαντικά προβλήματα πριν από 20 χρόνια και σήμερα θεωρούνται ασήμαντα. Ένας επιστομικός που επιρρεάστηκε αρκέτα είναι ο τομέας της Όρασης των Υπολογιστών (Computer Vision) και πιο συγκεκριμένα, το πρόβλημα της αναγνώρισης και εντοπισμού ανθρώπινης δράσης σε βίντεο.

1.1 Περιγραφή Προβλήματος

Η πρόκληση της αναγνώρισης και εντοπισμού ανθρώπινης δράσης έχει δύο κύριους στόχους:

1. Την αυτόματη αναγνώριση και ταξινόμησή οποιασδήποτε ανθρώπινης δραστηριότητας στο βίντεο.
2. Τον αυτόματο εντοπισμό αυτής της δράσης στο βίντεο

1.1.1 Αναγνώριση ανθρώπινης δραστηριότητας

Λαμβάνοντας υπόψη την αναγνώριση της ανθρώπινης δράσης, ένα βίντεο μπορεί να αποτελείται μόνο από ένα άτομο που κάνει κάτι. Ωστόσο, αυτό είναι ένα ιδανικό σενάριο. Στις περισσότερες περιπτώσεις, τα βίντεο περιέχουν πολλά άτομα, που εκτελούν πολλαπλές ενέργειες ή ενδέχεται να μην δρουν καθόλου σε ορισμένα τμήματα του βίντεο. Έτσι, ο στόχος μας δεν είναι μόνο να ταξινομήσουμε μια δράση, αλλά να αποκομίσουμε τα χρονικά όρια κάθε δράσης

1.1.2 Εντοπισμός ανθρώπινης δραστηριότητας

Παράλληλα με την αναγνώριση της ανθρώπινης δράσης, ένα άλλο πρόβλημα είναι να προσδιορίσουμε χωρικά όρια κάθε δράσης. Συνήθως, αυτό σημαίνει να καθορίσουμε ένα δισδιάστατο πλαίσιο οριοθέτησης για κάθε εικόνα βίντεο, το οποίο περιέχει τον δρόντα. Φυσικά, αυτό το κουτί οριοθέτησης κινείται μαζί με τον ηθοποιό.

1.2 Εφαρμογές

Το πεδίο της Ανθρώπινης Δράσης Αναγνώρισης και Εντοπισμού έχει πολλές εφαρμογές που περιλαμβάνουν ανάλυση περιεχομένου με βάση το περιεχόμενο, αυτοματοποιημένη κατάτμηση

βίντεο, συστήματα ασφάλειας και επιτήρησης, αλληλεπίδρασης ανθρώπου υπολογιστή. Η τεράστια διαθεσιμότητα δεδομένων (ειδικά των βίντεο) δημιουργεί την ανάγκη να βρεθούν τρόποι για να επωφεληθούμε απ΄ αυτά. Περίπου 2,5 δισεκατομμύρια εικόνες μεταφορτώνονται στη βάση δεδομένων του Facebook κάθε μήνα, περισσότερες από 34K ώρες βίντεο στο YouTube και περίπου 5K εικόνες κάθε λεπτό. Επιπλέον, υπάρχουν περίπου 30 εκατομμύρια κάμερες παρακολούθησης στις ΗΠΑ, πράγμα που σημαίνει περίπου 700 ώρες βίντεο ανά ημέρα. Όλα αυτά τα δεδομένα πρέπει να χωριστούν σε κατηγορίες ανάλογα με το περιεχόμενό τους προκειμένου να γίνουν πιο εύκολα προς αναζήτηση. Η διαδικασία αυτή γίνεται, συνήθως, χειρωνακτικά από έναν χρήστη που συνδέει το κάθε βίντεο με λέξεις-κλειδιά ή ετικέτες. Ωστόσο, οι περισσότεροι χρήστες αποφεύγουν να το κάνουν αυτό, τόσα πολλά βίντεο καταλήγουν χωρίς πληροφορίες σχετικά με τις ετικέτες. Αυτή η κατάσταση δημιουργεί την ανάγκη δημιουργίας αλγορίθμων για αυτοματοποιημένη εύρεση του κατάλληλου βίντεο με βάση το περιεχόμενο του.

Ένα άλλο πεδίο εφαρμογών είναι η περίληψη βίντεο. Αυτές οι εφαρμογές χρησιμοποιούνται συνήθως σε ταινίες ή αθλητικές εκδηλώσεις. Στις ταινίες, οι αλγόριθμοι ανάλυσης βίντεο μπορούν να δημιουργήσουν ένα μικρό βίντεο που περιέχει όλες τις σημαντικές στιγμές της ταινίας. Αυτό μπορεί να επιτευχθεί επιλέγοντας τμήματα βίντεο στα οποία λαμβάνει χώρα μια σημαντική ενέργεια, όπως η δολοφονία του κακοποιού της ταινίας. Στις αθλητικές εκδηλώσεις, οι εφαρμογές περίληψης βίντεο περιλαμβάνουν τη δημιουργία αυτόματων βίντεο προβολής, όπως π.χ. ένα βίντεο που περιέχει όλα τα επιτευχθέντα γκολ σ΄ έναν ποδοσφαιρικό αγώνα.

Επιπλέον, η αναγνώριση της ανθρώπινης δράσης μπορεί να αντικαταστήσει τους ανθρώπινους χειριστές στα συστήματα επιτήρησης. Μέχρι τώρα, τα συστήματα ασφαλείας περιλαμβάνουν ένα σύστημα πολλαπλών καμερών που τα χειρίζεται ένας άνθρωπος χειριστής, ο οποίος κρίνει εάν ένα άτομο ενεργεί κανονικά ή όχι. Τα συστήματα αυτόματης ταξινόμησης ενέργειας μπορούν να ενεργούν όπως ο άνθρωπος, και αμέσως να κρίνουν εάν υπάρχει κάποιους είδους περίεργη συμπεριφορά κρίνετε εάν υπάρχει ανωμαλία στον άνθρωπο.

Τελευταίο αλλά όχι ασήμαντο, ένα άλλο πεδίο εφαρμογής σχετίζεται με την αλληλεπίδραση ανθρώπου-υπολογιστή. Ρομποτικές εφαρμογές βοηθούν τους ηλικιωμένους να αντιμετωπίζουν τις καθημερινές τους ανάγκες. Επίσης, οι εφαρμογές παιχνιδιών που χρησιμοποιούν το Kinect δημιουργούν νέα επίπεδα εμπειρίας παιχνιδιού χωρίς την ανάγκη ενός ελεγκτή φυσικού παιχνιδιού.

1.3 Προκλήσεις και Datasets

Υπάρχουν διάφοροι τύποι ανθρώπινων δραστηριοτήτων. Ανάλογα με την πολυπλοκότητά τους, θεωρούμε ότι οι ανθρώπινες δραστηριότητες ταξινομούνται σε τέσσερις διαφορετικές κατηγορίες επίπεδα: χειρονομίες, ενέργειες, αλληλεπιδράσεις και δραστηριότητες ομάδας. Οι χειρονομίες είναι στοιχειώδεις κινήσεις του σώματος ενός ατόμου και είναι το ατομικό στοιχεία που περιγράφουν την ουσιαστική κίνηση ενός ατόμου. «Το στρίψιμο του βραχίονα» και «της ανύψωσης του ποδιού» είναι καλά παραδείγματα χειρονομιών. Οι ενέργειες είναι δραστηριότητες ενός ατόμου που μπορούν να αποτελούνται από πολλαπλές χειρονομίες που οργανώνονται προσωρινά, όπως «περπάτημα», «χαιρετισμός» και 'μπουνία'. Οι αλληλεπιδράσεις είναι ανθρώπινες δραστηριότητες που περιλαμβάνουν δύο ή περισσότερα άτομα και / ή αντικείμενα. Για παράδειγμα, «δύο άτομα που αγωνίζονται» είναι μια αλληλεπίδραση μεταξύ δύο ανθρώπων και «ενός ατόμου που κλέβει μια βαλίτσα από κάποιον άλλο» είναι μια αλληλεπίδραση ανθρώπου-αντικειμένου που περιλαμβάνει δύο ανθρώπους και ένα αντικείμενο. Τέλος, οι δραστηριότητες ομάδας είναι οι δραστηριότητες που εκτελούνται από εννοιολογικές ομάδες που αποτελούνται από πολλαπλά πρόσωπα και / ή αντικείμενα. «Μια ομάδα ανθρώπων που κάνουν πορεία», «μια ομάδα που έχει συνάντηση» και «δύο ομάδες που παίζουν ξύλο» είναι τυπικά παραδείγματα αυτών.

Η μεγάλη ποικιλία ανθρώπινων δραστηριοτήτων και εφαρμογών δημιουργεί πολλές προκλήσεις

που περιλαμβάνουν συστήματα αναγνώρισης της δράσης. Οι σημαντικότερες προκλήσεις περιλαμβάνουν μεγάλες διακυμάνσεις της εμφάνισης των ανθρώπων που δρουν, αλλαγές στην οπτική γωνία της κάμερας, αποκλείσεις, μη-άκαμπτες κινήσεις κάμερας κλπ. Επιπλέον, ένα μεγάλο πρόβλημα είναι ότι υπάρχουν πάρα πολλές κατηγορίες δράσης που σημαίνει ότι η χειροκίνητη συλλογή του δείγματος εκπαίδευσης είναι απαγορευτική. Επίσης, ορισμένες φορές, το λεξιλόγιο περιγραφής των δράσεων δεν είναι καλά καθορισμένο. Όπως δείχνει το σχήμα 1.1, η ενέργεια «Ανοίγω» μπορεί να περιλαμβάνει πολλά είδη ενεργειών, γι αύτό πρέπει προσεκτικά να αποφασίσουμε ποια έννοια αυτής της πράξης θα λάβουμε υπόψη.

Σχήμα 1.1: Παραδείγματα της δράσης «Ανοίγω»

Προκειμένου να αντιμετωπιστούν αυτές οι προκλήσεις, έχουν δημιουργηθεί διάφορα σύνολα δεδομένων για ανθρώπινες δράσεις, προκειμένου να αναπτυχθούν ισχυρά συστήματα αναγνώρισης της ανθρώπινης δράσης και αλγόριθμοι ανίχνευσης. Τα πρώτα σύνολα δεδομένων περιέλαμβαν έναν δρων χρησιμοποιώντας μιας στατικής κάμερα πάνω σε ομοιογενή υπόβαθρα. Παρόλο που αυτά τα σύνολα δεδομένων συνείσφεραν στο να σχεδιάσουμε τους πρώτους αλγορίθμους αναγνώρισης δράσης, δεν ήταν σε θέση να αντιμετωπίσουν αποτελεσματικά τις παραπάνω προκλήσεις. Έτσι λοιπόν οδηγηθήκαν στον να δημιουργήσουμε σύνολα δεδομένων που περιέχουν πιο αμφιλεγόμενα βίντεο, όπως το Joint-annotated Human Motion Database(JHMDB) (**Kuehne11**) και UCF-101 (**soomro2012ucf101**). Αυτά τα dataset περιλαμβάνουν μόνο ανθρώπινες ενέργειες, η δεύτερη κατηγορία που αναφέρθηκε πιο πριν.

1.3.1 JHMDB Dataset

Το σύνολο δεδομένων JHMDB (**Jhuang: ICCV: 2013**) είναι ένα πλήρες σχολιασμένο σύνολο δεδομένων για ανθρώπινες ενέργειες και ανθρώπινες πόζες. Αποτελείται από 21 κατηγορίες δράσεων και 928 κλιπ που εξάγονται από την βάση δεδομένων κίνησης του ανθρώπου (HMDB51) **Kuehne11**. Αυτό το σύνολο δεδομένων περιέχει κομμένα βίντεο με διάρκεια μεταξύ 15 έως 40 καρέ. Κάθε κλιπ σχολιάζεται για κάθε καρέ χρησιμοποιώντας μια δισδιάστατη στάση και περιέχει μόνο 1 ενέργεια. Προκειμένου να εκπαιδεύσουμε το μοντέλο μας για τον εντοπισμό των ενεργειών, τροποποιούμε της δισδιάστατες πόζες σε δισδιάσταστα πλαίσια που περιέχουν ολόκληρη τη στάση του δρώντα σε κάθε καρέ. Υπάρχουν διαθέσιμα 3 διαφορετικά χωρίσματα για να εκπαιδευτεί ένα μοντέλο, τα οποία προτείνουν οι συγγραφείς. Επιλέξαμε το πρώτο που περιέχει 660 βίντεο στο εκπαιδευτικό σετ και 268 για επικύρωση.

Μεχρι εδω εχω ελεγξει

1.3.2 UCF-101 Dataset

Το σύνολο δεδομένων UCF-101 (**soomro2012ucf101**) περιέχει 13320 βίντεο από 101 κατηγορίες δράσεων. Από αυτά, για 24 τάξεις και 3194 βίντεο χωροχρονικές σχολιασμοί

περιλαμβάνονται. Αυτό σημαίνει ότι υπάρχει ένα 2Δ οριοθετημένο πλαίσιο που περιβάλλει τον ηθοποιό για κάθε πλαίσιο στο οποίο λαμβάνει χώρα μια δράση. Τους διαχωρίζουμε σε 2284 βίντεο για εκπαιδευτικό σετ και 910 για δοκιμή επικύρωσης σύμφωνα με το αρχικά προτεινόμενο σχίσιμο. Για τα δεδομένα εκπαίδευσης, υπάρχουν βίντεο μέχρι 641 καρέ, ενώ σε στοιχεία επικύρωσης ο μέγιστος αριθμός πλαισίων είναι 900. Κάθε βίντεο, τόσο για την εκπαίδευση όσο και για την επικύρωση, δεν είναι έγκυρη, συμπεριλαμβανομένων μερικές φορές περισσότερες από 1 ενέργειες που πραγματοποιούνται ταυτόχρονα. Λάβαμε σχολιασμούς από **singh2016online** επειδή οι συγγραφείς που πρότειναν σχολιασμούς περιέχουν κάποια λάθη.

1.4 Μοτιατιον ανς ὃντιβυτιονς

Τα τρέχοντα επιτεύγματα στα δίκτυα αναγνώρισης αντικειμένων και στα δίκτυα 3Δ συνάθροισης για την αναγνώριση ενεργειών μας προκάλεσαν να δοκιμάσουμε για να τα συνδυάσουμε, προκειμένου να επιτύχουμε τα καλύτερα αποτελέσματα για τον εντοπισμό της δράσης. Εισάγουμε μια νέα δομή δικτύου εμπνευσμένη από το ϛιτε ΔΒΛΠ: ϑουρναλς / ϛορρ / Ηου"Σ17, ϛιτε ΔΒΛΠ: ϑουρναλς / ϛορρ / αβσ-1712-09184, ϛιτε Ρεν: 2015: ΦΡΤ: 2969239.2969250 από ϛιτε ϑθφαστερ2ρϛνν.

Οι συνεισφορές μας είναι οι εξής: 1) Δημιουργούμε ένα νέο πλαίσιο για τον εντοπισμό των ενεργειών που επεκτείνει τον κώδικα που λαμβάνεται από τα γρηγορότερα Ρ"ΝΝ, 2) Προσπαθούμε να δημιουργήσουμε ένα δίκτυο για την πρόταση ακολουθιών πλαισίων οριοθέτησης σε βίντεο κλιπ τα οποία μπορεί να περιέχουν μια ανάρτηση που εκμεταλλεύεται των χωροχρονικών χαρακτηριστικών που μας παρέχουν οι 3Δ ὃνολυτιονς, 3) δημιουργούμε έναν αλγόριθμο σύνδεσης για τη σύνδεση προτεινόμενων ακολουθιών οριοθετώντας κουτιά για να εξάγουμε σωλήνες υποψήφιας ενέργειας και 4) προσπαθούμε να βρούμε τους καταλληλότερους χάρτες χαρακτηριστικών για την ταξινόμησή τους.

1.5 Τηεσις στρυϛτυρε

Η υπόλοιπη διατριβή οργανώνεται ως εξής. Το κεφάλαιο 2 παρέχει μια γενική εισαγωγή στις τεχνικές μάθησης μηχανής που χρησιμοποιούνται σήμερα. Στη συνέχεια, παρουσιάζουμε τα βασικά στοιχεία των συστημάτων αναγνώρισης αντικειμένων και παράλληλα με τις λειτουργίες απώλειας και τις μετρήσεις αξιολόγησης συνηθίζαμε. Επίσης, το Κεφάλαιο 2 παρουσιάζει μια σύντομη επισκόπηση της βιβλιογραφίας σχετικά με την αναγνώριση και τον εντοπισμό της ανθρώπινης δράσης. Το Κεφάλαιο 3 εισάγει το πρώτο βασικό στοιχείο του δικτύου μας, Τυβε Προποσαλ Νετωορκ (ΤΠΝ), ένα δίκτυο που προτείνει Τυβες οφ Ιντερεστ (ΤΙς), οι οποίες είναι ακολουθίες των πλαισίων οριοθέτησης, με πιθανό να περιέχουν μια εκτελεσθείσα ενέργεια. Επιπλέον, περιέχει όλες τις προτεινόμενες αρχιτεκτονικές για την επίτευξη αυτού του στόχου. Το κεφάλαιο 4 προτείνει αλγόριθμους για τη σύνδεση των προτεινόμενων ΤΟΙ από κάθε τμήμα βίντεο και παρουσιάζονται οι επιδόσεις των προτάσεων. Στο Κεφάλαιο 5 παρουσιάζουμε όλες τις προσεγγίσεις ταξινόμησης που χρησιμοποιήσαμε για τον σχεδιασμό της αρχιτεκτονικής μας και ορισμένα αποτελέσματα ταξινόμησης. Το Κεφάλαιο 6 χρησιμοποιείται για συμπεράσματα, περίληψη της συμβολής μας μαζί με πιθανές μελλοντικές εργασίες.