

Κεφάλαιο 1

Επίλογος - Μελλοντικές επεκτάσεις

1.1 Επίλογος

Σε αυτή τη διατριβή εξερευνούμε το πρόβλημα της αναγνώρισης και του εντοπισμού της δράσης. Σχεδιάζουμε ένα δίκτυο βασισμένο στην προσέγγιση των **ΔΒΛΠ:θουρναλς/ζορρ/Ηου"Σ17** σε συνδυασμό με ορισμένα στοιχεία από τους **ΔΒΛΠ:θουρναλς/ζορρ/αβσ-1712-09184**, **Ρεν:2015:ΦΡΤ:2969239.2969250**, **Γιρσηικ:2015:ΦΡ:2919332.2920125**, **ΔΒΛΠ:θουρναλς/ζορρ/αβσ-1903-00304** και **ηαρα3δςννς**.

Γράψαμε μια πψτορση υλοποίηση παίρνοντας κώδικα μόνο από το **θθφαστερ2ρςνν**. Επιπλέον, γράψαμε τον δικό μας κώδικα χρησιμοποιώντας μερικές λειτουργίες της γλώσσας CUDA που έχουν σχεδιαστεί από εμάς (όπως υπολογισμός των βαθμολογιών σύνδεσης, τροποποίηση σωλήνων κλπ).

Προσπαθήσαμε να σχεδιάσουμε το TPN , ένα δίκτυο για την προτάσεις ToIs, ακολουθίες πλαισίων δηλαδή σε δεδομένο βίντεο, εμπνευσμένο από το RPN του Faster R-CNN. Το σχεδιάσαμε χρησιμοποιώντας γενικευμένα anchors και όχι συγκεκριμένα για κάθε σύνολο δεδομένων. Προσπαθούμε δηλαδή να γενικεύσουμε την προσέγγισή μας για διάφορα σύνολα δεδομένων, αντίθετα με την προσέγγιση που προτείνεται από το **ΔΒΛΠ:θουρναλς/ζορρ/αβσ-1712-09184**, στην οποία χρησιμοποιούνται τα πιο συχνά εμφανιζόμενα anchors για κάθε σύνολο δεδομένων.

Επιπροσθέτως, σχεδιάσαμε έναν αφελή αλγόριθμο σύνδεσης για τη σύνδεση των προτεινόμενων ToIs με βάση αυτόν που προτάθηκε από τους **ΔΒΛΠ:θουρναλς/ζορρ/αβσ-1712-09184**. Στην προσέγγισή μας, χρησιμοποιούμε την ίδια πολιτική βαθμολόγησης, η οποία είναι ένας συνδυασμός των βαθμολογιών της πιθανότητας δράσης και του σκορ επικάλυψης. Η κύρια διαφορά είναι ότι αποφεύγουμε να υπολογίζουμε πιθανούς συνδυασμούς χρησιμοποιώντας ένα όριο ενημέρωσης, που μπορεί να ανανεώνεται. Επίσης, δοκιμάσαμε κι άλλον έναν αλγόριθμο σύνδεσης εμπνευσμένος από τους **ΔΒΛΠ:θουρναλς/ζορρ/αβσ-1903-00304**. Ωστόσο, η εφαρμογή μας δεν

ήταν πολύ καλή, συνεπώς δεν εξερευνήσαμε όλες τις δυνατότητες του.

Τέλος, διερευνήσαμε αρκετούς ταξινομητές για το στάδιο ταξινόμησης του δικτύου. Αυτοί είναι: ένα RNN, ένας Γραμμικός ταξινομητής, ένα SVM και ένα MLP. Χρησιμοποιήσαμε μια εφαρμογή απ' το Φαστ Ρ³NN για τον ταξινομητή SVM, η οποία περιελάμβανε την διαδικασία εκπαίδευσης μέσω σκληρών αρνητικών. Εξετάσαμε μερικές τεχνικές εκπαίδευσης για βέλτιστη απόδοση ταξινόμησης και 2 εκπαιδευτικές προσεγγίσεις για τον ταξινομητή MLP, την κλασσική και μία που χρησιμοποιούμε προεξαγόμενα χαρακτηριστικά.

1.2 Μελλοντικές επεκτάσεις

Υπάρχουν πολλά περιθώρια βελτίωσης για το δίκτυό μας, προκειμένου να επιτευχθεί τελευταίας τεχνολογίας αποτελέσματα. Οι σημαντικότερες περιγράφονται στις επόμενες παραγράφους.

Βελτίωση των προτάσεων του TPN Υλοποιήσαμε 2 δίκτυα για την πρόταση ακολουθιών από πλαίσια σε ένα τμήμα βίντεο. Πετύχαμε περίπου 63% βαθμολογία recall για τη διάρκεια του δείγματος ίση με 16 καρέ και περίπου 80% recall για τη διάρκεια του δείγματος ίση με 8. Αυτά τα σκορ δείχνουν ότι υπάρχει αρκετός χώρος για βελτίωση ειδικά για την περίπτωση με δείγμα 16 καρέ. Παρόλο που έχουν διερευνηθεί πολλές αρχιτεκτονικές δικτύων για παλινδρόμηση, μια καλή ιδέα θα ήταν να δοκιμάσουμε άλλα δίκτυα, δεν είναι απαραίτητα εμπνευσμένη από δίκτυα εντοπισμού αντικειμένων όπως κάναμε εμείς. Επιπλέον, προσθέτοντας έναν παράγοντα λ στον τύπο του training loss θα ήταν μια καλή ιδέα και θα διερευνούσε ποια είναι η καλύτερη προσέγγιση αυτού. Έτσι, η απώλεια εκπαίδευσης θα μπορούσε να οριστεί ως:

$$L = \sum_i L_{cls}(p_i, p_i^*) + \lambda_1 \sum_i p_i^* L_{reg}(t_i, t_i^*) + \lambda_2 \sum_i q_i^* L_{reg}(c_i, c_i^*) \quad (1.1)$$

Επιπλέον, θα ήταν μια καλή ιδέα να χρησιμοποιήσουμε την μέθοδο που ΣΣΔ (ΔΒΛΠ:θουρναλς/σoρp/ΛιυΑΕΣΡ15) προτείνει RoIs αντί για το RPN, για να συγκρίνετε το αποτέλεσμα. Τέλος, θα μπορούσαμε να πειραματιστούμε χρησιμοποιώντας τα δίκτυα Feature Pyramid (8099589), τα οποία θα μπορούσαν να επεκταθούν σε 3 διαστάσεις ως ένα άλλο δίκτυο εξαγωγής χαρακτηριστικών ή κάποιο άλλο είδος 3D ResNet.

Αλλαγή του αλγορίθμου σύνδεσης Σε αυτή τη διατριβή, μια άλλη πρόκληση που αντιμετωπίσαμε ήταν η σύνδεση των προτεινόμενων ToIs για την πρόταση action tubes. Υλοποιήσαμε έναν πολύ αφελή αλγόριθμο, που δεν ήταν σε θέση να μπορεί να μας δώσει πολύ καλές προτάσεις παρά τις αλλαγές που προσπαθήσαμε να κάνουμε. Υλοποιήσαμε έναν άλλο αλγόριθμο σύνδεσης που ήταν βασισμένος στην εκτίμηση της χρονικής πρόοδο ενός action tube και την αλληλεπιδική του με άλλα. Αν και δεν μας έδωσε και πολύ καλές προτάσεις, πιστεύουμε ότι πρέπει να εξερευνήσουμε τις δυνατότητες αυτού του αλγορίθμου.

Κι αυτό επειδή είναι εκμεταλλεύεται την πρόοδο της ενέργειας, την οποία δεν είχε ο προηγούμενος αλγόριθμος.

Εξερεύνηση άλλων τεχνικών ταξινόμησης Για το στάδιο ταξινόμησης, πειραματιστήκαμε κυρίως πάνω σε έναν ταξινομητή SVM για το σύνολο δεδομένων JHMDB και δεν εμπλακήκαμε καθόλου με το σύνολο δεδομένων UCF. Ο πρώτος μας στόχος είναι να είμαστε σε θέση να εξάγουμε καλά αποτελέσματα ταξινόμησης για το σύνολο δεδομένων UCF. Πιστεύουμε ότι θα πρέπει να διερευνήσουμε τους χάρτες χαρακτηριστικών του UCF και τεχνικές που εφαρμόζονται στους χάρτες χαρακτηριστικών πριν από την ταξινόμηση. Επιπλέον, θα μπορούσαμε να δοκιμάσουμε άλλες τεχνικές ταξινόμησης όπως Random Forests ή να πειραματιστούμε περισσότερο με τον ταξινομητή RNN για το σύνολο δεδομένων UCF. Τέλος, μια άλλη διαδικασία ταξινόμησης θα ήταν μια καλή ιδέα, όπως η εξαγωγή πρώτα όλων των πιθανών action tubes και, στη συνέχεια, η χρήση άλλων δικτύων για εξαγωγή χαρακτηριστικών προκειμένου να ταξινομήσουμε τα action tubes.