

Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

Μεταπτυχιακή Διατριβή **Στην Ασφάλεια Υπολογιστών και Δικτύων**



Αναγνώριση Βίντεο, με Εφαρμογές στα Συστήματα Ασφαλείας

Ιωάννης Μέμος-Μπαγκράτσας

Επιβλέπων Καθηγητής
Νικόλαος Σκλάβος

Δεκέμβριος 2020

Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

Αναγνώριση Βίντεο, με Εφαρμογές στα Συστήματα Ασφαλείας

Ιωάννης Μέμος-Μπαγκράτσας

Επιβλέπων Καθηγητής
Νικόλαος Σκλάβος

Η παρούσα μεταπτυχιακή διατριβή υποβλήθηκε
προς μερική εκπλήρωση των απαιτήσεων για απόκτηση

μεταπτυχιακού τίτλου σπουδών
στην Ασφάλεια Υπολογιστών και Δικτύων

από τη Σχολή Θετικών και Εφαρμοσμένων Επιστημών
του Ανοικτού Πανεπιστημίου Κύπρου

Δεκέμβριος 2020

Περίληψη

Οι πρόσφατες εξελίξεις στα φορητά συστήματα καταγραφής βίντεο και η ευκολία πρόσβασης σε αυτά έχουν οδηγήσει στην αύξηση των βίντεο που μεταδίδονται καθημερινά στο διαδίκτυο. Αυτό έχει επηρεάσει σημαντικά το ρυθμό επέκτασης του διαδικτύου των πραγμάτων με την προσθήκη φθηνών συστημάτων ασφαλείας για την καταγραφή και μετάδοση βίντεο. Τέτοια συστήματα συχνά έχουν χαμηλές προδιαγραφές ασφαλείας παρά το γεγονός ότι χρησιμοποιούνται για να παρέχουν δικαστικά τεκμήρια εγκλήματος. Εγκληματίες μπορούν εύκολα να αποκτήσουν πρόσβαση, να κλέψουν και να πλαστογραφήσουν βίντεο στο σημείο όπου πολλές φορές να είναι αδύνατη η διάκριση ενός αυθεντικού βίντεο από ένα πλαστό. Επιπλέον, οι εξελίξεις στον τομέα της μηχανικής μάθησης έχουν οδηγήσει στην αύξηση ψηφιακά δημιουργημένων βίντεο γνωστών ως Deep Fakes με στόχο τον εξευτελισμό, τη δολιοφθορά, απειλή και προπαγάνδα σε κάθε τομέα της σύγχρονης ζωής, κοινωνικό, πολιτικό και στρατιωτικό. Όπως είναι αναμενόμενο, η ακεραιότητα ενός βίντεο δε μπορεί να θεωρηθεί δεδομένη και πρέπει σε πολλές περιπτώσεις να ελέγχεται. Η επιστήμη της Δικανικής Υπολογιστών έχει προτείνει αρκετές μεθόδους, ενεργητικές και παθητικές, για την προστασία και απόδειξη της ακεραιότητας των βίντεο στην πηγή ή κατά την απόκτηση. Ωστόσο, η ανάγκη για πιο ακέραιες και αποδοτικές μεθόδους αυξάνεται παράλληλα με το μέγεθος και τον αριθμό των βίντεο και καθώς οι αλγόριθμοι συμπίεσης γίνονται συνεχώς πιο αποδοτικοί. Στόχος της παρούσας μεταπτυχιακής διατριβής είναι να προτείνει μία νέα μέθοδο αναγνώρισης πλαστογραφίας, βασισμένη στα χαρακτηριστικά της Πυκνής Οπτικής Ροής. Η συγκεκριμένη μέθοδος εφαρμόζεται σε ένα Raspberry Pi 4B στοχεύοντας στην προσομοίωση ενός περιβάλλοντος διαδικτύου των πραγμάτων, με μία συσκευή χαμηλού κόστους, ικανή να αναγνωρίσει πλαστογραφία σε στατικά βίντεο συστημάτων ασφαλείας. Το τελικό εργαλείο ελέγχεται χρησιμοποιώντας δύο σύνολα πλαστών βίντεο και αποδεικνύει ότι η μέθοδος είναι ικανή να αναγνωρίσει πλαστογραφίες μέσω αντιγραφής με αναδιάταξη, εισαγωγής και διαγραφής ενώ είναι ανθεκτική σε μέτρια ποσοστά συμπίεσης και θορύβου. Ολοκληρώνοντας, η μεταπτυχιακή διατριβή προτείνει μία μεθοδολογία για την αυτοματοποίηση της ανίχνευσης όπως επίσης παραθέτει προτάσεις για μελλοντική έρευνα.

Λέξεις κλειδιά: Δικανική Υπολογιστών, βίντεο, πολυμέσα, αναγνώριση, πλαστογραφία, παραβίαση, ακεραιότητα, παθητικές μέθοδοι, ενεργητικές μέθοδοι, οπτική ροή, Gunnar Farneback, διαδίκτυο των πραγμάτων, IoT, OpenCV, Raspberry Pi.

Summary

The recent advances in mobile video recording systems and the ease of access to them have led to an increased amount of videos being streamed and uploaded on the internet daily. This has had a major impact on the expansion rate of the Internet of Things with the addition of affordable security systems that capture and stream video content. These systems regularly come with minimal security standards despite the fact that their goal is to provide court evidence. Criminals are often able to easily access, steal and manipulate videos to the point where it is many times impossible to tell the difference between an original and a forged one. Moreover, the advances in machine learning have led to an increase of computer generated videos known as deep fakes that are used for humiliation, sabotage, threat and propaganda in every aspect of modern life, social, political or militaristic. As expected, the integrity of a video cannot be taken for granted and should in many cases be evaluated. The scientific branch of Computer Forensics has proposed various methods, both active and passive, for protecting the confidentiality and integrity of videos either at the source or on acquisition. In spite of that, the need for more robust and effective methods rises as videos grow in size and number while compression algorithms are becoming increasingly more efficient. The goal of this master thesis is to propose a new method on forgery detection based on the characteristics of Dense Optical Flow. This method is applied on a Raspberry Pi 4B which serves to simulate an IoT environment with a low cost device that is capable of recognizing forgery in static CCTV captured videos. The tool is tested using two datasets of forged videos and proves that this method is effective in detecting copy-move, insertion and deletion forgeries while being robust against medium rates of compression and noise. Ultimately, this master thesis proposes a methodology on automating the process of identification as well as recommendations for future research.

Keywords: Computer Forensics, video, multimedia, recognition, forgery, tampering, integrity, passive methods, active methods, optical flow, Gunnar Farneback, internet of things, IoT, OpenCV, Raspberry Pi.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω από καρδιάς τον Καθηγητή μου, Δρ. Νικόλαο Σκλάβο για την ευκαιρία που μου έδωσε να εργαστώ πάνω σε ένα τόσο σημαντικό θέμα. Η καθοδήγηση και στήριξη που μου παρείχε στο διάστημα εκπόνησης υπήρξε ανεκτίμητη.

Επίσης, με την ολοκλήρωση των σπουδών μου, θα ήθελα να ευχαριστήσω την οικογένειά μου και κυρίως τη θεία μου, Χρυσούλα, που με παρότρυνε να ακολουθήσω το μεταπτυχιακό πρόγραμμα της Ασφάλειας Υπολογιστών και Δικτύων. Δεν ανέμενα ότι θα ανακάλυπτα το στόχο της ζωής μου. Τέλος, ευχαριστώ τη σύντροφό μου, Αθηνά, για την απaráμιλλη υπομονή και ενδυνάμωση όταν ένιωθα πολύ «λίγος» για το μέγεθος του παρόντος εγχειρήματος.

Σας ευχαριστώ όλους από καρδιάς.

Με εκτίμηση,
Ιωάννης Μέμος-Μπαγκράτσας

Περιεχόμενα

Κεφάλαιο 1.....	1
1.1 Στόχος της μεταπτυχιακής διατριβής.....	3
1.2 Δομή της μεταπτυχιακής διατριβής.....	3
Κεφάλαιο 2.....	5
2.1 Χαρακτηριστικά βίντεο - εισαγωγικά.....	5
2.2 Ενεργητικές μέθοδοι.....	6
2.2.1 Χρήση υδατογραφήματος.....	6
2.2.2 Χρήση του blockchain.....	8
2.3 Παθητικές μέθοδοι.....	9
2.3.1 Μέθοδοι βασισμένες σε αλλοιώσεις συμπίεσης.....	11
2.3.2 Μέθοδοι βασισμένες σε αλλοιώσεις εισαγωγής θορύβου.....	12
2.3.3 Μέθοδοι βασισμένες στην οπτική ροή.....	13
2.3.4 Υβριδικές μέθοδοι.....	14
2.3.5 Μηχανική μάθηση.....	15
2.4 Σχετικές έρευνες.....	16
2.5 Αλγόριθμος οπτικής ροής του Gunnar Farneback.....	20
2.6 Σύγκριση των μεθόδων Lucas-Kanade και Gunnar Farneback.....	21
2.7 Σημαντικότητα της έρευνας.....	22
Κεφάλαιο 3.....	23
3.1 Βασικά δομικά στοιχεία του εργαλείου.....	23
3.1.1 Raspberry Pi 4b.....	23
3.1.2 OpenCV και Python.....	25
3.1.3 Matplotlib.....	26
3.2 Υλοποίηση.....	26
Κεφάλαιο 4.....	31
4.1 Παράμετροι και σύνολα βίντεο.....	31
4.2 Πειράματα.....	33
4.2.1 Λείες μπάλες σε γραφείο.....	33
4.2.2 Λείες μπάλες σε τραχύ δάπεδο.....	36
4.2.3 Λείες μπάλες σε γραφείο με κίνηση πτώσης.....	38
4.2.4 Αυτοκίνητα διαφορετικού τύπου σε οριζόντια κίνηση.....	40
4.2.5 Μπάλες κλωστής σε λείο δάπεδο.....	42
4.2.6 Αυτοκίνητα ίδιου τύπου σε διαγώνια κίνηση.....	43
4.2.7 Αφαίρεση πλαισίου σε βίντεο χώρου στάθμευσης.....	45
4.2.8 Χώρος στάθμευσης με χαμηλή οπτική ροή.....	47
4.2.9 Χώρος στάθμευσης με αυξομείωση των τιμών φωτεινότητας.....	49

4.3 Έλεγχος αντοχής σε συμπίεση	51
4.4 Ανάλυση αποτελεσμάτων και βελτιστοποιήσεις.....	52
Κεφάλαιο 5.....	57
Βιβλιογραφία.....	59
Παράρτημα Α.....	63
Παράρτημα Β.....	67
B.1 Απαιτούμενο υλικό και συναρμολόγηση.....	67
B.2 Απαιτούμενο λογισμικό και εγκατάσταση.....	69
B.3 Σύνδεση μέσω SSH και VNC.....	73
B.4 Δημιουργία αρχείου pythοn και εκτέλεση	74

Κεφάλαιο 1

Εισαγωγή

Η τελευταία δεκαετία έχει επιφέρει ραγδαίες εξελίξεις τόσο στην ψηφιακή επεξεργασία αρχείων πολυμέσων όσο και στην πρόσβαση του μέσου ανθρώπου σε συστήματα ικανά για την επίτευξή της. Η χρήση πολυμέσων εικόνας και βίντεο αποτελεί πλέον σημαντικό παράγοντα στα μέσα κοινωνικής δικτύωσης, σε διαφημιστικές ή πολιτικές καμπάνιες, στη διασκέδαση, στην τηλεκπαίδευση και σε διάφορους ερευνητικούς τομείς όπως στατιστική, ασφάλεια, μηχανική μάθηση κ.α. Η τεχνολογική άνοδος σε συσκευές κινητής τηλεφωνίας κάθε κόστους (Sklavos & Toulιου, 2007) έχει οδηγήσει στη συχνή αντικατάσταση του συμβατικού κειμένου με αρχεία βίντεο και εικόνας. Συσκευές μικρού κόστους (Sklavos & Souras, 2006) και με συχνά περιορισμένες προδιαγραφές ασφαλείας (Sklavos, Chaves, Di Natale, & Regazzoni, 2017) όπως κλειστά κυκλώματα τηλεόρασης (CCTV) και κάμερες αυτοκινήτου χρησιμοποιούνται τακτικά για την απόκτηση βίντεο ως δικαστικά τεκμήρια. Επιπλέον, η επέκταση του διαδικτύου των πραγμάτων (Internet of Things - IoT) με τη χρήση συσκευών όπως ενσωματωμένων συστημάτων με FPGA (Pantopoulou & Sklavos, 2020), (Sklavos & Francese, 2020), (Antoniadis, Sklavos, & Kavum, 2020) καθώς επίσης μικροελεγκτών Arduino (Isa & Sklavos, 2017), (Panagiotou, Sklavos, & Zaharakis, 2018) και υπολογιστών Raspberry Pi, εξοπλισμένων με κάμερες, έχει αμβλύνει τον όγκο πολυμέσων που διακινούνται για προσωπική ή δημόσια χρήση. Οι δυνατότητες ασφαλείας που τέτοια συστήματα

παρέχουν περιορίζονται επιπλέον λόγω του μικρού μεγέθους τους και της δυσκολίας υλοποίησης λύσεων για την προστασία της ιδιωτικότητας σε επίπεδο υλικού (Limnaios, Sklavos, & Kouforanlou, 2019). Η συχνή παραβίαση τέτοιων συστημάτων σε συνδυασμό με την αύξηση της υπολογιστικής ισχύος των επεξεργαστών καθώς και βελτιστοποιήσεων σε επίπεδο εφαρμογών λογισμικού (Sklavos & Toulou, 2007) επιτρέπουν σε εύκολα προσβάσιμες φορητές συσκευές την ικανότητα υποκλοπής και κακόβουλης επεξεργασίας πολυμέσων υψηλής ανάλυσης.

Συμπερασματικά, αν και στην καθημερινότητα η επεξεργασία έχει ως κύριο στόχο ένα βελτιωμένο οπτικό αποτέλεσμα ή τη διευκόλυνση της ερευνητικής ανάλυσης (Sklavos & Francese, 2020), αλλοιώσεις μη ορατές από το ανθρώπινο μάτι συχνά εφαρμόζονται για την απόκρυψη στοιχείων ή την πλαστογράφιση πολυμέσων. Οι πρόσφατες εξελίξεις στη μηχανική μάθηση (machine learning) και βαθιά εκμάθηση (deep learning) οξύνουν το πρόβλημα της πλαστογράφησης καθιστώντας εφικτή τη δημιουργία ψεύτικων βίντεο (deep fakes) με στόχο την παραπληροφόρηση, σκευωρία, απειλή και εκβιασμό σε κοινωνικό, πολιτικό και στρατιωτικό επίπεδο.

Η ευκολία στην παραβίαση της αξιοπιστίας πολυμέσων έχει οδηγήσει στην ανάγκη εύρεσης στιβαρών μεθόδων για την απόδειξή της. Διάφορες λύσεις που έχουν προταθεί για την προστασία παραβίασης συστημάτων του IoT (Sklavos, Chaves, Di Natale, & Regazzoni, 2017), (Limnaios, Sklavos, & Kouforanlou, 2019) είναι ικανές να παρέχουν έμμεσα ένα ποσοστό ασφαλείας στα καταγεγραμμένα πολυμέσα από την μη εξουσιοδοτημένη απόκτησή τους, προστατεύοντας κυρίως τα συστήματα καταγραφής από μορφές παραβίασης (Panagiotou, Sklavos, & Zaharakis, 2018). Η ανάγκη για την επιμέρους ενίσχυση της ιδιωτικότητας και ακεραιότητας των πολυμέσων προκύπτει ως μία επιπλέον ζώνη προστασίας όταν τα συστήματα καταγραφής παραβιάζονται και καθώς η απόκτησή των πολυμέσων είναι επίσης εφικτή κατά τη διαδικτυακή μετάδοσή τους (Spanaki & Sklavos, 2018), (Sklavos, Hübner, Göhringer, & Kitsos, 2013, pp. 135-158) ή κατά την κλοπή υλικού. Η επιστήμη της δικανικής υπολογιστών έχει παρουσιάσει λύσεις στην εύρεση χειραγώγησης πολυμέσων, ωστόσο οι συνεχείς τεχνολογικές εξελίξεις των δυνατοτήτων επεξεργασίας (Sklavos & Toulou, 2007) δημιουργούν ανερχόμενες προκλήσεις που απαιτούν αντιμετώπιση. Ιδιαίτερη βαρύτητα αποκτά η ικανότητα υλοποίησης τέτοιων λύσεων στο περιβάλλον του IoT λόγω του μεγάλου όγκου διακινούμενων δεδομένων πολυμέσων που οφείλεται στην ταχύτατη επέκτασή του. Στο πλαίσιο αυτής της ανάγκης, η παρούσα διατριβή αρχικά ερευνά και

αναλύει σημαντικές μεθόδους απόδειξης πλαστογραφίας πολυμέσων και προτείνει μία νέα μεθοδολογία απόδειξης πλαστογραφίας βίντεο. Επεξηγείται το θεωρητικό υπόβαθρο και παρουσιάζεται ένα εργαλείο ικανό να αναγνωρίσει τρεις μορφές πλαστογραφίας και εφαρμόσιμο σε περιβάλλον κυκλωμάτων ασφαλείας και συσκευών του IoT.

1.1 Στόχος της μεταπτυχιακής διατριβής

Στόχος της παρούσας μεταπτυχιακής διατριβής είναι αρχικά η θεωρητική ανάλυση των σημαντικών ομάδων μεθόδων ανίχνευσης πλαστογραφίας σε πολυμέσα βίντεο. Ξεκινώντας, θα πραγματοποιηθεί μία βιβλιογραφική ανασκόπηση των ερευνών που έχουν ως στόχο τη σύγκριση των μεθόδων κάθε ομάδας ώστε να καθοριστεί η βάση για την έναρξη της έρευνας. Ακολούθως, θα μελετηθούν οι σχετικές μέθοδοι που έχουν προταθεί και θα αναλυθούν τα αποτελέσματα των αντίστοιχων υλοποιήσεων. Με τον καθορισμό της ερευνητικής βάσης και την κατανόηση των συγκεκριμένων ερευνών θα προταθεί και θα υλοποιηθεί μία νέα μέθοδος αναγνώρισης πλαστογραφίας σε πολυμέσα βίντεο. Σκοπός είναι η εφαρμογή της μεθόδου με τη δημιουργία ενός εργαλείου χρησιμοποιώντας τη βιβλιοθήκη OpenCV σε γλώσσα Python. Το εργαλείο θα υλοποιηθεί πειραματικά σε μία συσκευή Raspberry Pi 4B για την προσομοίωση ενός περιβάλλοντος διαδικτύου των πραγμάτων (IoT). Τέλος, μέσω των πειραμάτων εκτέλεσης το αποτέλεσμα της μεταπτυχιακής διατριβής θα αξιολογηθεί ως προς την αξιοπιστία και την απόδοσή του και θα διατυπωθούν προτάσεις για μελλοντική έρευνα με στόχο τη βελτιστοποίηση και την επέκτασή του.

1.2 Δομή της μεταπτυχιακής διατριβής

Η παρούσα μεταπτυχιακή διατριβή χωρίζεται σε 8 κεφάλαια. Αρχικά, το 2^ο κεφάλαιο αποτελεί μία βιβλιογραφική ανασκόπηση των μεθόδων πλαστογραφίας σε πολυμέσα βίντεο και της απόδειξης αυτής σε κάθε περίπτωση. Επιπλέον, αναφέρονται οι σύγχρονες προκλήσεις της δικανικής υπολογιστών και οι έρευνες που εμπίπτουν στο θεωρητικό υπόβαθρο του παραδοτέου εργαλείου. Συγχρόνως, επισημαίνονται οι λόγοι που καθιστούν σημαντική την παρούσα έρευνα. Το 3^ο κεφάλαιο περιλαμβάνει το σχεδιασμό και την επεξήγηση του εργαλείου της έρευνας. Περιγράφεται η μεθοδολογία που ακολουθήθηκε και αναλύονται με σαφήνεια τα βήματα του αλγορίθμου.

Ακολουθως, στο 4^ο κεφάλαιο παρουσιάζονται σενάρια εκτέλεσης του εργαλείου με τη χρήση έτοιμων συνόλων από πλαστά βίντεο. Κάθε σενάριο αναλύεται σε σύγκριση με το αυθεντικό βίντεο προς απόδειξη της αξιοπιστίας του εργαλείου. Ως αποτέλεσμα, στο 5^ο κεφάλαιο αναφέρονται στο σύνολό τους τα πορίσματα της μεταπτυχιακής διατριβής και παρατίθενται προτάσεις για μελλοντική έρευνα. Το 6^ο κεφάλαιο περιλαμβάνει τις βιβλιογραφικές αναφορές που αποτέλεσαν την ερευνητική βάση της διατριβής ενώ το 7^ο κεφάλαιο αποτελεί παράρτημα για την παράθεση του κώδικα του εργαλείου. Τέλος το 8^ο κεφάλαιο αποτελεί ένα επιπλέον παράρτημα με αναλυτικές οδηγίες για τη συναρμολόγηση του υλικού και την εγκατάσταση του λογισμικού για το εργαλείο της μεταπτυχιακής διατριβής.

Κεφάλαιο 2

Θεωρητικό Υπόβαθρο

Το παρόν κεφάλαιο αποτελεί τη βιβλιογραφική ανασκόπηση της μεταπτυχιακής διατριβής. Γίνεται κατανομή των μορφών πλαστογραφίας βίντεο και αντίστοιχα της ανίχνευσής τους. Περιγράφεται το θεωρητικό υπόβαθρο για τον καθορισμό του ερευνητικού άξονα της μεταπτυχιακής διατριβής.

2.1 Χαρακτηριστικά βίντεο - εισαγωγικά

Για την κατανόηση των επόμενων εννοιών είναι απαραίτητο να εξηγηθούν κάποια χαρακτηριστικά πολυμέσων βίντεο. Οι διάφορες λύσεις που έχουν προταθεί βασίζονται σε αυτά τα χαρακτηριστικά ώστε να αποδείξουν την ακεραιότητα των πολυμέσων. Η θεωρητική βάση είναι πως αλλοιώσεις που εφαρμόζονται σε ένα βίντεο οδηγούν σε αποκλίσεις στη συσχέτιση αυτών των χαρακτηριστικών και επομένως η κατανόησή τους είναι ύψιστης σημασίας. Ένα βίντεο αποτελεί την αποτύπωση φυσικών περιβαλλόντων και αντικειμένων μέσω του φακού της βιντεοκάμερας σε συνεχή πλαίσια. Κάθε πλαίσιο υφίσταται αρχικά επεξεργασία και στη συνέχεια κωδικοποιείται σε ένα αποθηκευτικό μέσο. Κάθε στάδιο από τη λήψη έως την κωδικοποίηση εισάγει συγκεκριμένα χαρακτηριστικά θορύβου ή παραμόρφωσης που εξαρτώνται τόσο από

την κάμερα όσο και από το λογισμικό επεξεργασίας που χρησιμοποιεί. Η κωδικοποίηση εισάγει επιπλέον χαρακτηριστικά όπως πρόβλεψη κίνησης και θόρυβο. Στις επόμενες δύο υποενότητες περιγράφονται οι δύο μεγάλες κατηγορίες προστασίας ακεραιότητας ή απόδειξης παραβίασης σε βίντεο. Αναφορικά, οι ενεργητικές μέθοδοι χρησιμοποιούνται για την οικειοποίηση βίντεο στην πηγή, δηλαδή τη χρήση κρυπτογραφίας, στεγανογραφίας και Blockchaining και για την προστασία της ακεραιότητας του πολυμέσου. Η δεύτερη κατηγορία περιλαμβάνει τις παθητικές μεθόδους οι οποίες στοχεύουν στην ανακάλυψη παραβίασης της ακεραιότητας των πολυμέσων. Η μεταπτυχιακή διατριβή εστιάζει περισσότερο στις παθητικές μεθόδους οι οποίες και αναλύονται εκτενέστερα. Αν και οι ενεργητικές μέθοδοι εξακολουθούν να χρησιμοποιούνται σε συστήματα ζωντανής μετάδοσης και καταγραφής, ο όγκος των βίντεο που διακινούνται στο διαδίκτυο δημιουργεί την πρόκληση εύρεσης της αυθεντικής πηγής. Οι παθητικές μέθοδοι δεν εξαρτώνται από το αυθεντικό βίντεο ή τον/την κάτοχό του και επομένως αποκτούν μεγαλύτερη σημασία στη σύγχρονη εποχή.

2.2 Ενεργητικές μέθοδοι

Όπως έχει αναφερθεί, οι ενεργητικές μέθοδοι έχουν ως στόχο την προστασία της ακεραιότητας του βίντεο. Για την επίτευξή της στις περισσότερες περιπτώσεις εφαρμόζονται τεχνικές κρυπτογραφίας, στεγανογραφίας και χρήσης του blockchain. Οι παρακάτω κατηγορίες περιγράφονται συνοπτικά ενώ η βιβλιογραφία που τις στηρίζει αποτελεί ανασκόπηση ερευνών σε κάθε κατηγορία αντίστοιχα. Στόχος είναι να αναφερθούν στη γενικότητά τους τα συστατικά χαρακτηριστικά τέτοιων μεθόδων και να συγκριθούν με τις παθητικές ως προς την αποτελεσματικότητα και εφαρμοστικότητα τους.

2.2.1 Χρήση υδατογραφήματος

Οι (Shojanazeri, Wan Adnan, & Mumtadzah Syed Ahmad, 2013) περιγράφουν τα χαρακτηριστικά και τις κύριες τεχνικές υδατογράφησης για την απόδειξη υποκλοπής και παραβίασης του βίντεο. Αναφέρεται πως σε αντίθεση με πολυμέσα εικόνας, η παραβίαση βίντεο οδηγεί σε αλλοίωση του υδατογραφήματος και επισημαίνεται η ανάγκη για υδατογράφηση που διέπεται από στιβαρότητα και ανθεκτικότητα σε αλλοιώσεις. Η εφαρμογή υδατογραφήματος εισάγει ένα ποσοστό θορύβου το οποίο

πρέπει να περιορίζεται στο ελάχιστο δυνατό. Παράλληλα, ένα σύστημα υδατογράφησης πρέπει να είναι ικανό να αναγνωρίσει το υδατογράφημα παρά τον επιπλέον θόρυβο που εισάγεται κατά τη μετάδοση του βίντεο ή σε περίπτωση της παραβίασής του. Σημαντικοί παράγοντες που πρέπει να λαμβάνονται υπόψη είναι η ταχύτητα εφαρμογής και εξαγωγής του υδατογραφήματος, η χρήση ισχυρών κρυπτογραφικών κλειδιών για την προστασία της υπογραφής και η ανθεκτικότητα στην ανακάλυψη της υδατογράφησης μέσω στατιστικών μεθόδων. Επιπλέον, πρέπει να περιορίζεται το ποσοστό λάθους στο ελάχιστο ώστε το σύστημα να καθίσταται αξιόπιστο στην εξαγωγή του υδατογραφήματος. Οι μέθοδοι υδατογράφησης σε χωρικό επίπεδο (spatial domain) εφαρμόζονται συνήθως με τη χρήση της τεχνικής του λιγότερο σημαντικού bit (Least Significant Bit - LSB). Μέθοδοι τέτοιου είδους παρέχουν ανθεκτικότητα στην περικοπή του βίντεο και καθιστούν την ανακάλυψη του υδατογραφήματος δύσκολη, αλλά είναι επιρρεπείς σε αλλοιώσεις από την εισαγωγή επιπλέον θορύβου και συμπίεσης. Μία επιπλέον κατηγορία υδατογράφησης σε χωρικό επίπεδο είναι οι τεχνικές βασισμένες στη συσχέτιση των τιμών χρώματος και φωτεινότητας των εικονοστοιχείων. Ένα ποσοστό ψευδοτυχαίου θορύβου εφαρμόζεται ομοιόμορφα στην εικόνα με τη χρήση κλειδιού το οποίο είναι απαραίτητο για τη μετέπειτα εξαγωγή του υδατογραφήματος. Αποτέλεσμα αυτού είναι το υδατογράφημα να εξάγεται με μεγαλύτερη ακρίβεια. Η δεύτερη μεγάλη κατηγορία μεθόδων υδατογράφησης χαρακτηρίζεται από την εισαγωγή της υπογραφής στο επίπεδο συχνοτήτων (frequency domain). Γίνεται ευρεία χρήση διακριτών μετασχηματισμών που συνήθως είναι μετασχηματισμοί συνημίτονου (Discrete Cosine Transformation - DCT), wavelets (Discrete Wavelet Transformation - DWT) και Φουριέ (Discrete Fourier Transformation - DFT). Ο εκάστοτε μετασχηματισμός παρέχει ένα πεδίο συχνοτήτων που χρησιμοποιείται για την παραγωγή ή/και εισαγωγή του υδατογραφήματος. Ο μετασχηματισμός DCT συνήθως εφαρμόζεται σε μη συμπιεσμένα αρχεία και παρέχει ισχυρή ανθεκτικότητα σε απλές μορφές επεξεργασίας όπως συμπίεση MPEG-x ή αλλαγές χρωμάτων και φωτεινότητας, ωστόσο είναι αδύναμος σε περιπτώσεις περικοπής, περιστροφής ή αυξομείωσης του βίντεο. Ο μετασχηματισμός DFT παρέχει μία απεικόνιση του πεδίου συχνοτήτων αλλά το αποτέλεσμα είναι η απώλεια χρόνου. Επομένως είναι χρήσιμος για να παρουσιάσει μία απεικόνιση από τις χρωματικές υποζώνες των πλαισίων ενός βίντεο ώστε να γίνει επιλογή αυτών για την εισαγωγή του υδατογραφήματος. Για την εξαγωγή του εφαρμόζεται αντίστροφος DFT. Το πλεονέκτημα χρήσης του DFT είναι η αντοχή στην αφαίρεση εικονοστοιχείων. Βασικό

του μειονέκτημα είναι ότι η εφαρμογή του σε κάθε πλαίσιο απαιτεί αυξημένη υπολογιστική ισχύ σε σχέση με τον DCT και η έλλειψη συσχέτισης με το αρχικό πλαίσιο, δηλαδή η επιλογή της σωστής υποζώνης για το υδατογράφημα είναι δύσκολη ώστε να μην είναι ορατή η αλλαγή στο συμπιεσμένο βίντεο. Η χρήση μετασχηματισμού DWT επιλύει αυτό το πρόβλημα καθώς παρουσιάζει ομοιότητες με τον τρόπο λειτουργίας του ανθρώπινου συστήματος όρασης (Human Visual System - HVS). Είναι εφικτή η επιλογή των υποζώνων που είναι λιγότερο αντιληπτές από το ανθρώπινο μάτι ώστε ο θόρυβος από την εισαγωγή του υδατογραφήματος να είναι ο ελάχιστος δυνατός. Επιπλέον δεν υπάρχει απώλεια χρόνου όπως με τον DFT και σχετικές έρευνες έχουν αποδείξει πως το αποτέλεσμα χαρακτηρίζεται από διαφάνεια και ισχυρή προστασία του υδατογραφήματος, έχοντας το μειονέκτημα υψηλών απαιτήσεων επεξεργασίας (Jeong, Moon, & Kim, 2008). Συνοπτικά, οι (Shojanazeri, Wan Adnan, & Mumtadzah Syed Ahmad, 2013) καταλήγουν πως η χρήση μετασχηματισμών στο επίπεδο συχνοτήτων είναι εξαιρετικά απαιτητική σε υπολογιστικούς πόρους ώστε να είναι εφαρμόσιμη σε πραγματικό χρόνο λήψης βίντεο. Η χρήση μεμονωμένων μετασχηματισμών προστατεύει την ακεραιότητα του βίντεο από συγκεκριμένες παραβιάσεις. Προτείνεται η παράλληλη χρήση μηχανικής μάθησης ώστε σε κάθε βίντεο να εφαρμόζεται ο κατάλληλος μετασχηματισμός όπως επίσης συστήνεται η χρήση υδατογράφησης στο ηχητικό σήμα του βίντεο.

2.2.2 Χρήση του blockchain

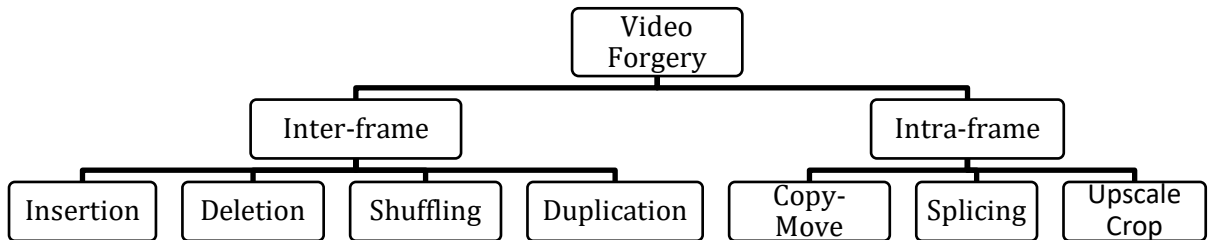
Μία δεύτερη κατηγορία ενεργητικών μεθόδων εφαρμόζει τη χρήση της δομής δεδομένων blockchain (Theodorou & Sklavos, 2019, p. 25). Κάθε κόμβος του blockchain περιέχει το hash του προηγούμενου κόμβου, ένα πεδίο με τη χρονική σήμανση της εισαγωγής του (timestamp) και τη ρίζα ενός δέντρου Merkle για τα δεδομένα του κόμβου (Kalis & Belloum, 2018). Κάθε κόμβος επιβεβαιώνει τον επόμενο στη δομή και επιβεβαιώνεται από τον προηγούμενο. Επομένως η διαδικασία της κρυπτογραφίας αλυσίδας παρέχει πολύ υψηλή ασφάλεια στην ακεραιότητα των δεδομένων κάθε κόμβου καθώς κάποιος για να παραλλάξει τα δεδομένα σε ένα κόμβο πρέπει πρώτα να κάνει το ίδιο για όλους τους προηγούμενους στην αλυσίδα (Theodorou & Sklavos, 2019, p. 26). Μία έρευνα που βασίστηκε στο blockchain (Yatskiv, Yatskiv, & Bandrivskyi, 2019) χρησιμοποιεί το εργαλείο ffmpeg για το διαχωρισμό του βίντεο σε πλαίσια (I, P και B) εισάγοντας σε μία δομή blockchain κάθε πλαίσιο. Η χρονική σήμανση

υποδηλώνει τη θέση του πλαισίου στο βίντεο. Το μειονέκτημα του αλγορίθμου είναι πως η ασφάλεια του βίντεο εξαρτάται άμεσα από την ασφάλεια του κόμβου ρίζας. Βασικό πλεονέκτημα είναι πως καθίσταται αδύνατο για κάποιον να ανακτήσει συγκεκριμένα πλαίσια του βίντεο αν δε σπάσει την κρυπτογράφηση των κόμβων για όλα τα προηγούμενα πλαίσια της αλυσίδας αντίστοιχα. Γενικά, οι συγκεκριμένες μέθοδοι βασίζονται σε ισχυρούς κρυπτογραφικούς αλγορίθμους και επομένως απαιτούν μεγάλη υπολογιστική ισχύ (Theodorou & Sklavos, 2019, p. 27) καθιστώντας δύσκολη την εφαρμογή τους σε συσκευές του διαδικτύου των πραγμάτων όπως στην περίπτωση της παρούσας έρευνας. Επιπλέον, η χρήση του blockchain παρουσιάζει το ίδιο πρόβλημα με το γενικό σύνολο των ενεργητικών μεθόδων. Δηλαδή μόνο το αυθεντικό βίντεο μπορεί να επιβεβαιώσει την πλαστογραφία και επομένως η παροχή του για ανάλυση βρίσκεται στη δικαιοδοσία του κατόχου. Αν ο κάτοχος χάσει για τον οποιοδήποτε λόγο την πρόσβαση σε αυτό τότε η μέθοδος δημιουργεί μεγαλύτερο πρόβλημα από αυτό που επιλύει.

2.3 Παθητικές μέθοδοι

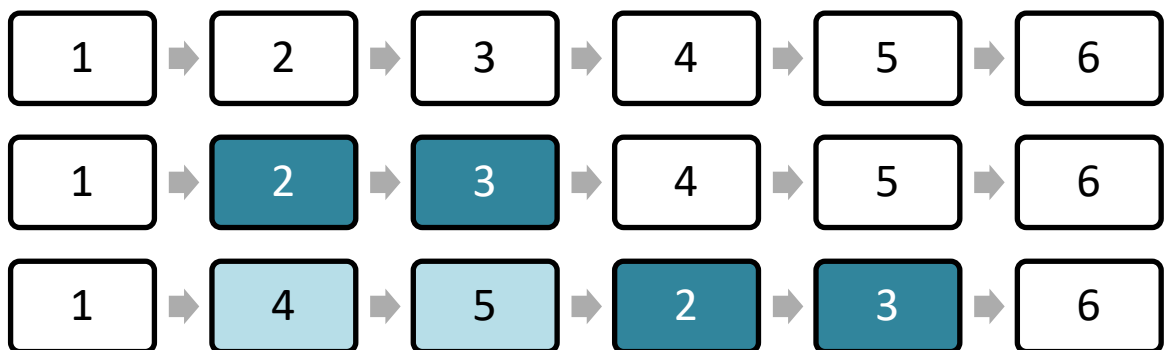
Το πρόβλημα χρήσης ενεργητικών μεθόδων είναι πως η ακεραιότητα ενός βίντεο δε μπορεί να εξακριβωθεί χωρίς την ύπαρξη του αυθεντικού. Η πλειοψηφία των βίντεο που μεταφορτώνονται στα μέσα κοινωνικής δικτύωσης όπως Facebook, Instagram, Youtube, TikTok κ.α. δε λαμβάνει πάντα επεξεργασία για την προστασία της ακεραιότητάς τους καθώς ο μέσος χρήστης συχνά δε γνωρίζει για αυτές. Ταυτόχρονα, τέτοια επεξεργασία απαιτεί χρόνο που πολλοί χρήστες δεν προθυμοποιούνται να αφιερώσουν. Σε πολλές περιπτώσεις συστημάτων ασφαλείας τέτοιες μέθοδοι δεν εφαρμόζονται καθώς τα μέσα κάλυψης δεν παρέχουν πάντα την απαιτούμενη υπολογιστική ισχύ. Η επιστήμη της δικανικής υπολογιστών έχει προτείνει μεθόδους οι οποίες αντιμετωπίζουν το παραπάνω πρόβλημα και βασίζονται στην αποκλειστική χρήση του ύποπτου βίντεο. Αυτές οι μέθοδοι ονομάζονται παθητικές (passive methods) γιατί στηρίζονται στα χαρακτηριστικά των πολυμέσων βίντεο κατά τη λήψη τους και τις διαφοροποιήσεις ή ανωμαλίες που παρουσιάζουν μετά από επεξεργασία. Είναι γνωστό πως ένα βίντεο αποτελείται από μία διαδοχή πλαισίων (frames) και την κωδικοποίησή τους σε ένα ενιαίο αρχείο ώστε να δημιουργείται η φυσική αίσθηση κίνησης. Κάθε μορφή αλλοίωσης εφαρμόζεται πάνω σε ομάδες πλαισίων και βάσει αυτού διακρίνονται δύο κύριες κατηγορίες παραβίασης (Kaur & Kushwaha, 2018). Η

πρώτη κατηγορία περιλαμβάνει την επεξεργασία μεταξύ πλαισίων (inter frame) και συγκεκριμένα την αντικατάσταση, αντιγραφή ή διαγραφή ομάδων από πλαίσια σε χρονικό επίπεδο (temporal domain). Η δεύτερη κατηγορία περιλαμβάνει την επεξεργασία εντός πλαισίου (intra frame), δηλαδή τη διαχείριση συγκεκριμένων αντικειμένων μέσω περικοπής (crop), αυξομείωσης (rescale), επανάληψης και ανακατανομής (copy-move) ή διαγραφής (delete) εντός ομάδων πλαισίων στο χωρικό επίπεδο (spatial domain). Συνοπτικά αυτές οι ομάδες παρουσιάζονται στο Σχήμα 1.



Σχήμα 1. Τεχνικές πλαστογραφίας βίντεο

Οι μέθοδοι πλαστογραφίας μεταξύ πλαισίων ή αλλιώς χρονικές παραβιάσεις (temporal tampering) έχουν ως στόχο την απόκρυψη, επανάληψη ή αντικατάσταση συμβάντων και έχουν περιορισμένες δυνατότητες ως προς την ποικιλομορφία του αποτελέσματος. Το Σχήμα 2 παρουσιάζει ένα παράδειγμα πλαστογραφίας μεταξύ πλαισίων.



Σχήμα 2. Παράδειγμα πλαστογραφίας εντός πλαισίων με αναδιάταξη των πλαισίων 2 και 3.

Η παραβίαση εντός πλαισίων ή αλλιώς χωρική παραβίαση (spatial tampering) δημιουργεί περισσότερες προκλήσεις καθώς μπορεί να εφαρμοστεί σε συγκεκριμένες περιοχές εντός των πλαισίων. Στόχος τέτοιων παραβιάσεων είναι η απόκρυψη αντικειμένων ή τροποποίηση συμβάντων και όχι τόσο η ανακατανομή τους στο χρόνο. Ένα παράδειγμα τέτοιας παραβίασης παρουσιάζεται στην Εικόνα 1.



Εικόνα 1. Πλαστογραφία εντός περιοχών πλαισίου

2.3.1 Μέθοδοι βασισμένες σε αλλοιώσεις συμπίεσης

Με τον όρο «αλλοίωση συμπίεσης» αναφερόμαστε σε σφάλματα που προκύπτουν κατά τη συμπίεση ενός ήδη αποκωδικοποιημένου βίντεο και είναι ανεξάρτητα από τη μέθοδο πλαστογραφίας. Οι παθητικές μέθοδοι ανίχνευσης συχνά αναφέρονται στην εύρεση διπλής συμπίεσης καθώς το ήδη συμπιεσμένο κατά τη λήψη βίντεο αποκωδικοποιείται, επεξεργάζεται και πλαστογραφείται ώστε να κωδικοποιηθεί εκ νέου στο τελικό βίντεο (Shelke & Kasana, 2020, p. 11). Αλλοιώσεις προκύπτουν λόγω του τρόπου λειτουργίας των εκδόσεων της ευρέως χρησιμοποιημένης μεθόδου MPEG (εκδόσεις 1, 2, 3, 4). Η χωρική συμπίεση MPEG γίνεται σε 3 στάδια και έχει ως αποτέλεσμα την παραγωγή τριών διαφορετικών τύπων καρέ/πλαισίων. Αυτοί οι τύποι πλαισίων συνιστούν ένα σύνολο πλαισίων με κοινά χαρακτηριστικά γνωστό ως GOP (Group of Frames). Ένα βίντεο που έχει υποστεί συμπίεση MPEG αποτελείται από πολλά GOP και τα πλαίσια που συνιστούν το κάθε GOP παρουσιάζουν συσχετίσεις μεταξύ τους. Ο διαχωρισμός των πλαισίων ενός GOP γίνεται σε I, P και B πλαίσια. Τα I (intracoded) πλαίσια απαιτούν το μέγιστο αποθηκευτικό χώρο καθώς συμπιέζονται λιγότερο από τα υπόλοιπα με συμπίεση JPEG και περιλαμβάνουν περισσότερες πληροφορίες ορίζοντας την αρχή ενός GOP. Συγχρόνως, η συμπίεσή τους είναι ανεξάρτητη από τη συμπίεση των υπόλοιπων πλαισίων του συνόλου. Τα P (predicted) πλαίσια εξαρτώνται από τα I πλαίσια καθώς περιέχουν μόνο πληροφορίες για τις αλλαγές σε σχέση με αυτά και

συμπιέζονται περισσότερο. Τα B (bi-directional) πλαίσια εξαρτώνται από τα I και P πλαίσια μεταξύ των οποίων παρεμβάλλονται για την εκτίμηση της κίνησης αλλά δε σχετίζονται με άλλα B πλαίσια. Λαμβάνουν έντονη συμπίεση και αποκωδικοποιούνται τελευταία εντός του GOP στο οποίο ανήκουν. Η εισαγωγή αλλοιώσεων προκύπτει όταν η πλαστογραφία εφαρμόζεται σε συγκεκριμένα πλαίσια ενός GOP και όχι στο σύνολό του. Για παράδειγμα αν μεταξύ δύο GOP αφαιρεθεί το I πλαίσιο ή υποστεί εσωτερική επεξεργασία τότε το επόμενο B ή P πλαίσιο δε θα μπορεί να συσχετίσει τις περιοχές όπου η πληροφορία εξαρτάται από το συγκεκριμένο I πλαίσιο και το βίντεο θα φαίνεται παραμορφωμένο. Όμοια αποτελέσματα παρατηρούνται κατά τη διαγραφή P πλαισίων που οδηγούν σε αλλοιώσεις στα εξαρτώμενα B πλαίσια. Οι μέθοδοι που στοχεύουν στην ανίχνευση τέτοιων αλλοιώσεων (Shelke & Kasana, 2020, p. 12) μελετούν τις συσχετίσεις συντελεστών μετασχηματισμών DCT μεταξύ των πλαισίων, τα χαρακτηριστικά ομοιογένειας μέσω του ιστογράμματος προσανατολισμένων διαβαθμίσεων (Histogram of Oriented Gradients - HOG) και συχνά κάνουν χρήση μοντέλων μηχανών διανυσμάτων υποστήριξης (Support Vector Machine - SVM) για την για την εκπαίδευση ταξινομητών. Ένα πρόβλημα τέτοιων μεθόδων είναι πως η εφαρμογή τους προϋποθέτει τη διαγραφή ή εισαγωγή συγκεκριμένων πλαισίων σε ένα GOP και είναι αδύνατο να ανιχνεύσουν την παραβίαση αν έχει χρησιμοποιηθεί ολόκληρο το GOP (Shelke & Kasana, 2020, p. 54). Ένα δεύτερο πρόβλημα είναι πως δεν παρέχουν ικανοποιητικά αποτελέσματα όταν το μέγεθος των GOP ενός βίντεο δεν είναι σταθερό και η εφαρμογή τους είναι ανέφικτη σε μη συμπιεσμένα βίντεο. Τέλος, εξαρτώνται πλήρως από την ύπαρξη συμπίεσης MPEG-x και δεν είναι αποτελεσματικές σε βίντεο συμπιεσμένα με H.264/AVC (Advanced Video Coding) καθώς η διαδικασία συμπίεσης είναι διαφορετική και παράγει χαμηλά bitrates (Johnston & Elyan, 2019).

2.3.2 Μέθοδοι βασισμένες σε αλλοιώσεις εισαγωγής θορύβου

Οι συγκεκριμένες μέθοδοι στηρίζονται σε χαρακτηριστικά θορύβου που εισάγονται από τις ψηφιακές βιντεοκάμερες και έχουν παρατηρηθεί να ακολουθούν συγκεκριμένα μοτίβα εντός των βίντεο (Shelke & Kasana, 2020, p. 15). Τέτοια χαρακτηριστικά είναι ο θόρυβος λήψης (Photon shot Noise - PSN), ο θόρυβος μοτίβου (Fixed Patter Noise - FPN), ο θόρυβος αισθητήρων (Sensor Pattern Noise - SPN) και ο θόρυβος μη ομοιομορφίας απόκρισης φωτογραφίας (Photo Response Non-Uniformity - PRNU). Η ανίχνευση συχνά απαιτεί την εξαγωγή των χαρακτηριστικών θορύβου και τη συσχέτισή

τους ανά τα πλαίσια του βίντεο. Μία βασική παραδοχή (Fayyaz, Anjum, Ziauddin, & Khan, 2020, p. 4) είναι πως η συσχέτιση μεταξύ υπολειμμάτων θορύβου σε διαδοχικά πλαίσια παρουσιάζει σημαντική απόκλιση σε περίπτωση πλαστογραφίας. Οι συγκεκριμένες μέθοδοι είναι αξιόπιστες και εξαιρετικά αποτελεσματικές όταν εφαρμόζονται σε μη συμπιεσμένα βίντεο με στατικό φόντο. Η συμπίεση μπορεί να οδηγήσει σε εισαγωγή επιπλέον θορύβου ή καταστροφή της συσχέτισης των χαρακτηριστικών θορύβου που εισάγονται από την κάμερα (Shelke & Kasana, 2020, pp. 15,54). Κατά δεύτερον, τα μοτίβα θορύβου έχουν μεταβληθεί με την πάροδο των χρόνων και την εξέλιξη των συσκευών καταγραφής και απαιτείται πλέον ο υπολογισμός των πιο πρόσφατων μοτίβων. Ένα ακόμα μειονέκτημα τέτοιων τεχνικών είναι πως τύπος θορύβου στο βίντεο επηρεάζει σε σημαντικό βαθμό την απόδοσή τους.

2.3.3 Μέθοδοι βασισμένες στην οπτική ροή

Η συγκεκριμένη κατηγορία περιλαμβάνει μεθόδους που βασίζονται σε χαρακτηριστικά της κίνησης όπως αυτά υπολογίζονται σε γειτονικά πλαίσια ενός βίντεο (Shelke & Kasana, 2020, p. 20). Τέτοια χαρακτηριστικά είναι η οπτική ροή, το υπόλειμμα κίνησης και τα υπολείμματα αντιστάθμισης κίνησης (Motion Compensated Edge Artifacts - MCEA). Τα υπολείμματα αντιστάθμισης κίνησης παρατηρούνται σε βίντεο που έχουν κωδικοποιηθεί με αλγορίθμους πρόβλεψης κίνησης (block motion compensation) και έχουν υποστεί διαγραφή πλαισίων. Η πρόβλεψη κίνησης ορίζει το κάθε πλαίσιο να αποκωδικοποιείται με βάση τα προηγούμενά του. Σε τέτοιες περιπτώσεις η διαγραφή πλαισίων οδηγεί στην εμφάνιση υπολειμμάτων εντός των επόμενων πλαισίων και μπορεί να προδώσει την ύπαρξη πλαστογραφίας. Η οπτική ροή (optical flow) μπορεί να αποτελέσει σημαντικό στοιχείο εύρεσης πλαστογραφίας καθώς σε ένα αυθεντικό βίντεο διατηρείται ομαλή. Διαγραφές ή εισαγωγές αντικειμένων σε χωρικό και χρονικό επίπεδο τείνουν να σπάνε την οπτική ροή και να δημιουργούν απότομες αυξομειώσεις στις τιμές της σε γειτονικά πλαίσια. Ένα βασικό μειονέκτημα των μεθόδων που στηρίζονται στα MCEA είναι η εξάρτησή τους από τα GOP ενός βίντεο. Οι περιορισμοί είναι παρόμοιοι με αυτούς των μεθόδων συμπίεσης καθώς τέτοιες μέθοδοι δεν είναι αξιόπιστες όταν οι παραβιάσεις έχουν γίνει σε ολόκληρα GOP, τα GOP έχουν διαφορετικά μήκη ή το βίντεο έχει συμπιεστεί με H264/AVC. Επιπλέον, όπως και η οπτική ροή, είναι μη αποδοτικές σε βίντεο με κινούμενο φόντο. Τέλος, η οπτική ροή αν και είναι ανθεκτική σε συμπίεση και εξαρτάται λίγο από το μήκος των GOP θεωρείται

αναξιόπιστη σε βίντεο με περιορισμένη κίνηση ή νυχτερινή λήψη και απαιτεί κάποια επιπλέον μέθοδο για τμηματοποίηση των πλαισίων όταν υπάρχουν πολλά κινούμενα αντικείμενα στο βίντεο. Οι στατιστικές μέθοδοι παρουσιάζουν κοινά σημεία με την οπτική ροή καθώς ελέγχουν σε γειτονικά πλαίσια τη συσχέτιση χαρακτηριστικών όπως υφή, φωτεινότητα, σκίαση, θόλωση, συντελεστές μετατροπών, περιγραφείς σημείων (SURF, SIFT, ORB, BRISK) και δυαδικά μοτίβα (Shelke & Kasana, 2020, p. 28). Παρουσιάζουν όμοια προβλήματα με τις μεθόδους ελέγχου κίνησης και συνήθως είναι περιορισμένες στην εύρεση μίας μορφής πλαστογραφίας. Συχνά απαιτείται ο συνδυασμός με άλλες μεθόδους ανίχνευσης για τη μείωση του χρόνου εκτέλεσης και την αύξηση της αποτελεσματικότητάς τους.

2.3.4 Υβριδικές μέθοδοι

Όπως αναφέρεται στα πορίσματα της έρευνας (Shelke & Kasana, 2020), οι περισσότερες έρευνες που έχουν διεξαχθεί στα πλαίσια της δικανικής πολυμέσων παρουσιάζουν αδυναμίες και συνήθως ανιχνεύουν μία ή δύο περιπτώσεις πλαστογραφίας. Για τη βελτιστοποίησή τους πολλές φορές είναι απαραίτητη η χρήση επιμέρους τεχνικών ανίχνευσης ώστε να αυξηθεί η ακρίβειά τους και η ανθεκτικότητα σε συμπίεση και θόρυβο. Τεχνικές που περιλαμβάνουν τη χρήση ασύνδετων μεταξύ τους χαρακτηριστικών τα οποία όμως μπορούν να ομαδοποιηθούν βάσει των ελέγχων που πραγματοποιούν ονομάζονται υβριδικές. Ένα παράδειγμα είναι η ανάλυση της οπτικής ροής όταν υπάρχει κίνηση πολλαπλών αντικειμένων σε ένα βίντεο. Μία μέθοδος που μπορεί να συμβάλει θετικά στην αξιοπιστία των αποτελεσμάτων της είναι η τμηματοποίηση της εικόνας σε υπερ-εικονοστοιχεία (superpixels) με τη μέθοδο SLIC (Simple Linear Iterative Clustering) (Pun & Chung, 2018). Σημαντικό χαρακτηριστικό των υβριδικών μεθόδων ανίχνευσης είναι η ανάγκη για πολύ μεγάλη υπολογιστική ισχύ. Η εύρεση σημείων και ανιχνευτών (keypoints and detectors) μέσω των αλγορίθμων SURF, SIFT, BRISK, ORB είναι μία συχνά εφαρμοσμένη μέθοδος σε πολυμέσα εικόνας και συνήθως απαιτεί μία βάση δεδομένων από εικόνες για σύγκριση. Στην περίπτωση τμηματοποίησης σε υπερ-εικονοστοιχεία για την εύρεση κοινών σημείων εντός της εικόνας χωρίς σύνολο αναφοράς εισάγεται πρόσθετο υπολογιστικό βάρος. Οι (Zhao, Wang, & Lu, 2018) χρησιμοποίησαν τη μέθοδο εξαγωγής σημαντικών σημείων με τον αλγόριθμο SURF (Speeded up Robust Features) σε συνδυασμό με τη βιβλιοθήκη εύρεσης κοντινότερων γειτόνων FLANN (Fast Library for Approximate Nearest Neighbors) για

την ανίχνευση πλαστογραφίας στο χρονικό επίπεδο χρησιμοποιώντας αντιστοίχιση χρωματικών τιμών του μοντέλου HSV το οποίο αναλύεται στο κεφάλαιο 3. Η μέθοδος τους αν και παρουσιάζει μεγάλη ακρίβεια στην εύρεση διαγραφής και εισαγωγής πλαισίων καθίσταται μη εφαρμόσιμη όταν υπάρχουν αλλαγές στο φόντο ή όταν η φωτογραφία για την εύρεση χαρακτηριστικών ομοιογένειας δεν έχει ληφθεί υπό τις ιδανικές συνθήκες. Αξίζει να αναφερθεί πως στα πλαίσια της παρούσας έρευνας η παραπάνω διαδικασία εφαρμόστηκε πειραματικά σε συσκευή Raspberry Pi 4B με παράλληλη τμηματοποίηση των πλαισίων για την εύρεση περιοχών ομοιογένειας στο χωρικό επίπεδο. Το αποτέλεσμα έδειξε πως για την ολοκλήρωση ανάλυσης σε βίντεο μικρού μεγέθους (350x250 pixels) και διάρκειας 15" απαιτήθηκαν κατά μέσο όρο 20' και επομένως η μέθοδος κρίθηκε μη εφαρμόσιμη.

2.3.5 Μηχανική μάθηση

Η μηχανική μάθηση (Machine Learning - ML) και η βαθιά εκμάθηση (Deep Learning - DL) έχουν χρησιμοποιηθεί σε μικρότερο βαθμό στη δικανική πολυμέσων. Βασικό πρόβλημα τέτοιων μεθόδων είναι οι απαιτήσεις εξαιρετικά μεγάλου αριθμού δεδομένων για την εξαγωγή των χαρακτηριστικών που έχουν προαναφερθεί και η ανάγκη για υψηλή επεξεργαστική ισχύ για την εξαγωγή τους και την εκπαίδευση των μοντέλων (Shelke & Kasana, 2020, p. 42). Μία συχνή τεχνική είναι η χρήση μοντέλων SVM για την εκπαίδευση ταξινομητών ώστε να είναι ικανοί να διακρίνουν πλαστά και αυθεντικά σημεία σε βίντεο. Το ιστόγραμμα προσανατολισμένων διαβαθμίσεων (HOG) χρησιμοποιείται επίσης σε μεγάλο εύρος για την ανίχνευση αντικειμένων βάσει των χαρακτηριστικών τους. Τα Συνελκτικά και Βαθιά Συνελκτικά Νευρωνικά Δίκτυα (Convolutional and Deep Convolutional Neural Networks - CNN, DCNN) εφαρμόζονται συχνά σε εφαρμογές αναγνώρισης πολυμέσων εικόνας και βίντεο λόγω της ικανότητάς τους να μαθαίνουν και να αναγνωρίζουν ομοιογένειες στα χαρακτηριστικά δυσδιάστατων εισόδων. Οι (Shelke & Kasana, 2020, pp. 42-44) έχουν παρουσιάσει επαρκώς μία συνολική έκθεση των πρόσφατων μεθόδων ανίχνευσης πλαστογραφίας με τη χρήση ML και DL. Συμπερασματικά έχουν καταλήξει ότι οι υπάρχουσες μέθοδοι παρουσιάζουν κατά μέσο όρο ελλείψεις ως προς τη δυνατότητα να αναγνωρίσουν τα πλαίσια όπου εφαρμόστηκε η πλαστογραφία (localization) και απαιτούν μεγάλο χρόνο εκτέλεσης και υπολογιστικής επεξεργασίας. Αρκετές μέθοδοι βασίζονται στη χρήση των I, P και B πλαισίων του GOP και όπως αναφέρθηκε, παρουσιάζουν μικρή αξιοπιστία

όταν τα GOP σε ένα βίντεο δεν έχουν σταθερό μήκος, όταν έχουν αφαιρεθεί πλήρως ή όταν δεν έχει εφαρμοστεί συμπίεση MPEG-x. Ωστόσο, οι συγγραφείς ενθαρρύνουν τη χρήση ML και DL καθώς τέτοιες μέθοδοι είναι εξαιρετικά πρόσφατες και υπάρχει μεγάλο ερευνητικό φάσμα που παραμένει αναξιοποίητο. Παράλληλα, η συνεχής αύξηση του αριθμού των Deep Fakes καθιστά τη μηχανική και βαθιά εκμάθηση απαραίτητες στην ανίχνευσή τους.

2.4 Σχετικές έρευνες

Η συγκεκριμένη υποενότητα αποτελεί τη βιβλιογραφική ανασκόπηση ερευνών που έχουν συσχετιστεί με την αναγνώριση παραβίασης βίντεο μέσω των χαρακτηριστικών της οπτικής ροής. Στα πλαίσια της μεταπτυχιακής διατριβής αποτελεί το μεταίχμιο μεταξύ του συνολικού ερευνητικού ορίζοντα της δικανικής πολυμέσων βίντεο και της παρούσας έρευνας ώστε να καταστήσει κατανοητά τα αποτελέσματα σχετικών ερευνών και τη σημαντικότητα της συγκεκριμένης.

Οι έρευνες που αναλύονται βασίζονται στη χρήση της μεθόδου υπολογισμού της οπτικής ροής Lucas-Kanade (Lucas & Kanade, 1981). Ξεκινώντας, η εξίσωση της οπτικής ροής εφαρμόζεται σε κάθε εικονοστοιχείο και είναι η ακόλουθη.

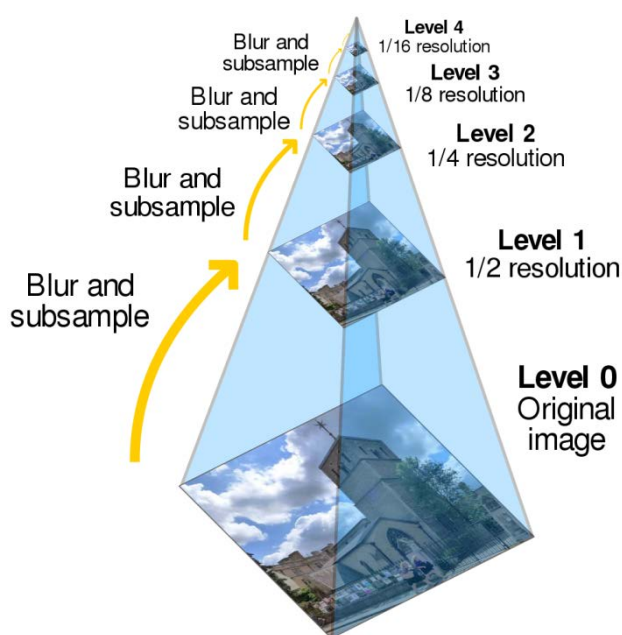
$$I_x u + I_y v + I_t = 0$$

Με βάση την παραπάνω εξίσωση, η μέθοδος Lucas-Kanade υπολογίζει το διάνυσμα κίνησης για ομάδες εικονοστοιχείων. Αρχικά η εικόνα διαχωρίζεται σε ομάδες από πίνακες γειτονικών εικονοστοιχείων π.χ. 3x3, 9x9 και στη συνέχεια γίνεται ο υπολογισμός της οπτικής ροής με τη μέθοδο των σταθμισμένων ελάχιστων τετραγώνων. Η εξίσωση είναι η ακόλουθη.

$$E_{LK} = \min \sum_{x \in \Omega} W^2(x) (I_x u + I_y v + I_t)^2$$

Για κάθε ομάδα υπολογίζεται η φωτεινότητα του κέντρου και των γειτόνων του και δεδομένου ότι ίδιες τιμές παρουσιάζονται μετατοπισμένες σε περιοχές του επόμενου πλαισίου (σταθερότητα φωτεινών περιοχών) γίνεται πρόβλεψη της μετατόπισης φωτεινότητας δηλαδή της κίνησης ενός σημείου του πλαισίου. Το αποτέλεσμα της

εφαρμογής του αλγορίθμου παράγει ένα σύνολο από πίνακες μετατόπισης διανύσματος που ορίζουν την οπτική ροή. Βασικό πλεονέκτημα της μεθόδου είναι πως η γραμμικότητά της επιτρέπει γρήγορους χρόνους εκτέλεσης. Μειονεκτήματά της είναι πως απαιτείται η ομαλότητα μετατόπισης των περιοχών φωτεινότητας για το κέντρο κάθε ομάδας εικονοστοιχείων και τους γείτονές του και ταυτόχρονα οι τιμές φωτεινότητας να παραμένουν σταθερές κατά τη μετατόπισή τους. Δηλαδή, η μέθοδος είναι πιο αποδοτική όταν τα αντικείμενα έχουν μικρή κίνηση στον τρισδιάστατο χώρο. Για την επίλυση του συγκεκριμένου προβλήματος προτείνεται παράλληλα η χρήση πυραμίδας Gauss (Εικόνα 2) της οποίας η συνεισφορά περιγράφεται παρακάτω.



Εικόνα 2. Μοντέλο της πυραμίδας Gauss

Μία υλοποίηση που έχει προταθεί (Chao, Jiang, & Tanfeng, 2012) βασίζεται στην ευαισθησία της οπτικής ροής σε πλαστογραφία χρονικού επιπέδου. Η μέθοδος που προτείνεται περιλαμβάνει τέσσερα στάδια υλοποίησης. Στο πρώτο στάδιο χρησιμοποιούνται δύο διαδοχικά πλαίσια και εφαρμόζεται σε καθένα μείωση του μεγέθους για τον περιορισμό των υπολογιστικών απαιτήσεων. Επιπλέον σε κάθε πλαίσιο εφαρμόζονται πέντε διαδοχικές σμικρύνσεις (κλιμάκωση πυραμίδας Gauss) ώστε να υπολογιστούν για κάθε μία τα διανύσματα μετατόπισης με τη μέθοδο Lucas-Kanade. Για κάθε πλαίσιο και ξεκινώντας αντίστοιχα από τη μικρότερη σμίκρυνση τα διανύσματα μετατόπισης προστίθενται στον πίνακα οπτικής ροής της αμέσως επόμενης βαθμίδας ώστε να αποφεύγεται το ενδεχόμενο ακύρωσης της ροής σε απότομες

μετατοπίσεις των φωτεινών περιοχών. Πρακτικά, σε περιπτώσεις όπου η κίνηση αντικειμένων είναι πολύ γρήγορη και οι μετατοπίσεις της φωτεινότητας ξεπερνάνε τα διαδοχικά εικονοστοιχεία η χρήση της πυραμίδας Gauss λειτουργεί για την πρόβλεψη της οπτικής ροής εντός αυτών. Στη συνέχεια, οι τιμές της οπτικής ροής για κάθε εικονοστοιχείο υπολογίζονται και παράγονται δύο ιστογράμματα για τη ροή στους άξονες X και Y. Η οπτικοποίηση του αποτελέσματος αποδεικνύει πως η εισαγωγή ή διαγραφή πλαισίων οδηγεί αντίστοιχα σε διακυμάνσεις αύξησης ή μείωσης της οπτικής ροής. Το δεύτερο στάδιο της υλοποίησης είναι η εύρεση ύποπτων περιοχών όπου η οπτική ροή αυξομειώνεται απότομα σε σχέση με ένα ορισμένο όριο (threshold). Για την περίπτωση εισαγωγής πλαισίων το βίντεο χωρίζεται σε τμήματα των 16 πλαισίων και η οπτική ροή υπολογίζεται μεταξύ του πρώτου και τελευταίου πλαισίου για κάθε τμήμα. Το όριο που καθορίζεται είναι η μέση οπτική ροή για όλο το βίντεο και υπολογίζεται για τους άξονες X και Y ξεχωριστά. Αν σε κάποιο παράθυρο η οπτική ροή για τους άξονες X ή Y είναι μεγαλύτερη από το διπλάσιο της μέσης οπτικής ροής του βίντεο στους δύο άξονες αντίστοιχα, τότε το παράθυρο θεωρείται ύποπτο. Για την εύρεση πλαστογραφίας πραγματοποιείται δυαδική αναζήτηση στα ύποπτα τμήματα με το κάθε τμήμα διαδοχικά να υφίσταται διαίρεση σε ίσα υπο-τμήματα. Για κάθε υπο-τμήμα εφαρμόζεται ξανά η παραπάνω διαδικασία διατηρώντας το ήδη καθορισμένο όριο και περιορίζοντας την αναζήτηση σε διαίρεση του αριστερού ή δεξιού υπο-τμήματος ανάλογα με το αν και ποιο υπο-τμήμα (αν όχι και τα δύο) παρουσιάζει οπτική ροή μεγαλύτερη του ορίου. Η δυαδική αναζήτηση ολοκληρώνεται όταν το μέγεθος ενός από τα δύο υπο-τμήματα γίνεται ίσο με 1 δηλαδή περιορίζεται στο πλαίσιο όπου έγινε η πλαστογραφία. Για την ανίχνευση αφαίρεσης πλαισίων η διαδικασία εφαρμόζεται σε ανά δύο γειτονικά πλαίσια και ακολουθεί παρόμοιο μοτίβο. Κατά τους δημιουργούς η μεθοδολογία μπορεί να αποδείξει τόσο την ύπαρξη προσθήκης ή αφαίρεσης πλαισίων σε βίντεο όσο και το σημείο όπου έχει γίνει η παραβίαση. Με βάση τα σύνολα πειραμάτων τους αναφέρουν πως στην εισαγωγή πλαισίων έχουν 95% επιτυχία ανακάλυψης με 98% ακρίβεια ενώ στην αφαίρεση 85% επιτυχία με 89% ακρίβεια.

Μία δεύτερη υλοποίηση (Wang, Wang, Jiang, & Sun, 2013) που χρησιμοποιεί τη μέθοδο Lucas-Kanade σε συνδυασμό με την παραγωγή μοντέλου πυραμίδας Gauss έχει βασιστεί στην διατήρηση της ομαλότητας της οπτικής ροής σε ένα στατικό βίντεο. Παρουσιάζει ομοιότητες με τη μέθοδο των (Chao, Jiang, & Tanfeng, 2012) καθώς υπολογίζει το μέγεθος (magnitude) της οπτικής ροής για τη στατιστική ανάλυση. Η θεωρητική βάση είναι πως οι απότομες αυξομειώσεις της οπτικής ροής όταν γίνεται εισαγωγή ή

αφαίρεση πλαισίων. Επιπλέον, με την παρατήρηση ότι η κατανομή πιθανότητας για τις παραλλαγές στην οπτική ροή ακολουθεί την κατανομή Gauss χρησιμοποιούνται στατιστικές μέθοδοι εύρεσης σημείων που δεν ακολουθούν αυτή την κατανομή για την απόδειξη πλαστογραφίας. Τα αποτελέσματα της μεθόδου τους στην εύρεση εισαγωγών, διαγραφών και αντιγραφών πλαισίων σε δύο διαφορετικά βίντεο παρουσιάζουν υψηλή ακρίβεια και το ποσοστό ψευδώς θετικών αποτελεσμάτων (false-positives) είναι 10%. Επιπροσθέτως, η μέθοδος κρίθηκε αρκετά ανθεκτική σε απωλεστική (lossy) συμπίεση.

Συνεχίζοντας, οι (Bidokhti & Ghaemmaghami, 2015) βασίστηκαν στην υλοποίηση των (Wang, Wang, Jiang, & Sun, 2013) εισάγοντας παράλληλα μία μέθοδο που περιλαμβάνει την επιλογή ύποπτων περιοχών στο χωρικό και χρονικό επίπεδο. Η μέθοδος στοχεύει στη βελτιστοποίηση του αλγορίθμου των (Wang, Wang, Jiang, & Sun, 2013) καθώς υπολογίζει την οπτική ροή ξεχωριστά για ύποπτες περιοχές εντός πλαισίων (Regions of Interest - ROI) και τις περιοχές εκτός αυτών. Ως ύποπτες ορίζονται οι περιοχές που δεν είναι στάσιμες και μπορούν να διαχωριστούν με τη χρήση γεωμετρικών σχημάτων. Οι συγγραφείς βασίζονται στην ιδέα ότι παραβιάσεις που επηρεάζουν την οπτική ροή μπορούν να εφαρμοστούν εντός των πλαισίων και όχι μόνο μεταξύ πλαισίων. Συγκεκριμένα, χρησιμοποιώντας το δείκτη $a(k)$ προκύπτει η παρατήρηση ότι πλαίσια που έχουν υποστεί περικοπή και εισαγωγή τμημάτων από άλλα πλαίσια του βίντεο παρουσιάζουν διαφορά στο δείκτη $a(k)$ σε σχέση με αυτά που έχουν παραμείνει ακέραια. Βασική προϋπόθεση για να συμβεί αυτό είναι τα τμήματα που έχουν μεταφερθεί σε περικομμένες περιοχές πλαισίων να μην αποτελούν πολλαπλάσια του συγκεκριμένου GOP. Κάτι τέτοιο εισάγει θόρυβο κατά την κωδικοποίηση με MPEG και αποτελεί τη βάση για την αποτελεσματικότητα της μεθόδου. Οι ύποπτες περιοχές διαχωρίζονται με τη χρήση μάσκας και το $a(k)$ υπολογίζεται για κάθε περιοχή, σημαντική και μη. Στους νέους δείκτες a_{ROI} και a_{other} εφαρμόζεται φίλτρο υπολογισμού του μέσου και η ανίχνευση πλαστογραφίας γίνεται με την εύρεση δευτερευόντων απότομων αυξομειώσεων του μεγέθους της ροής. Γίνεται αναφορά σε αυτές ως δευτερεύοντες καθώς σε κάθε GOP εμφανίζεται μία πρώτη αυξομείωση στο αρχικό πλαίσιο. Στη σύγκριση με τη μέθοδο των (Wang, Wang, Jiang, & Sun, 2013) η συγκεκριμένη υλοποίηση κρίνεται πιο αποδοτική από τους δημιουργούς της όταν η πλαστογραφία έχει εφαρμοστεί σε χρονικό και ταυτόχρονα χωρικό επίπεδο καθώς η πρώτη υποθέτει πως η πλαστογραφία μπορεί να προκύψει μόνο με τη διαγραφή, εισαγωγή ή αντιγραφή ολόκληρων πλαισίων και όχι περιοχών σε αυτά. Ένα μειονέκτημα της μεθόδου είναι ότι το αποτέλεσμα εξαρτάται από τη σωστή εφαρμογή

της μάσκας για το διαχωρισμό των περιοχών ενδιαφέροντος. Ταυτόχρονα, η μέθοδος είναι αναξιόπιστη αν δεν ικανοποιείται η βασική συνθήκη που λαμβάνει ως παράγοντα το μήκος του GOP στο οποίο εφαρμόζεται η αντιγραφή.

2.5 Αλγόριθμος οπτικής ροής του Gunnar Farneback

Ο Gunnar Farneback εισήγαγε τον αλγόριθμο υπολογισμού πυκνής οπτικής ροής ή διαφορετικά, το αλγόριθμο υπολογισμού της πολυωνυμικής επέκτασης των γειτονικών εικονοστοιχείων (Farneback, 2003). Συγκεκριμένα γίνεται η χρήση τετραγωνικών πολυωνύμων όπως φαίνεται παρακάτω

$$f(x) \sim x^T A x + b^T x + c$$

Ο A είναι συμμετρικός πίνακας, ο b είναι ένα διάνυσμα κατεύθυνσης και ο c είναι ένας πίνακας βαθμωτού μεγέθους (scalar). Το $f(x)$ αποτελεί την προσέγγιση των γειτονικών εικονοστοιχείων για το πλαίσιο A . Επομένως αν θεωρηθούν δύο πλαίσια A_1 και A_2 τότε η αρχική προσέγγιση είναι η εξής

$$\text{Αν } f_1(x) = x^T A_1 + b_1^T x + c_1 \text{ τότε υπάρχει}$$

$$f_2(x) = x^T A_2 + b_2^T x + c_2 \text{ τέτοιο ώστε}$$

$$A_2 = A_1 \text{ και } b_2 = b_1 - 2A_1 d$$

Επομένως, με τη μεταβλητή d να είναι το πεδίο επέκτασης στα γειτονικά εικονοστοιχεία ισχύει ότι η οπτική ροή για τα δύο πλαίσια είναι:

$$d = -\frac{1}{2} A_1^{-1} (b_2 - b_1)$$

Ένα πρόβλημα της αρχικής παραδοχής είναι πως τα πολυώνυμα είναι ίδια και το μόνο που αλλάζει είναι η επέκταση d . Όπως αναφέρεται (Farneback, 2003, p. 5) κάτι τέτοιο δεν ισχύει στην πραγματικότητα καθώς θα υπάρχει χωρική διαφορά μεταξύ των δύο πολυωνύμων. Το λάθος που εισάγεται είναι μικρό και μπορεί να θεωρηθεί αμελητέο όταν η επέκταση d παίρνει μικρές τιμές. Για την αποφυγή προβλημάτων όταν οι τιμές της d αυξάνονται προτείνεται η εισαγωγή μίας αρχική τιμής $\tilde{d}(x)$ τέτοια ώστε για δύο

πολύωνυμα στις θέσεις x και \tilde{x} να ισχύει $\tilde{x} = x + \tilde{d}(x)$. Δίνεται έτσι η δυνατότητα να υπολογιστεί το επιτρεπτό μέγεθος της μετατόπισης.

2.6 Σύγκριση των μεθόδων Lucas-Kanade και Gunnar Farneback

Η σύγκριση των δύο μεθόδων (De Boer & Kalksma, 2015) έχει πραγματοποιηθεί για την επιλογή της καταλληλότερης στον υπολογισμό μετατόπισης μη επανδρωμένου αεροσκάφους (Unmanned Aerial Vehicle - UAV). Τα πειράματα πραγματοποιήθηκαν σε ένα επεξεργαστή Intel i5-4300U χρησιμοποιώντας τις υλοποιήσεις των δύο αλγορίθμων μέσω της βιβλιοθήκη OpenCV . Τα αποτελέσματα δείχνουν σχετική ομοιότητα των δύο αλγορίθμων σε απόδοση όταν δεν εφαρμόζεται ομαδοποίηση των εικονοστοιχείων σε κάθε πλαίσιο. Στην περίπτωση εφαρμογής ελέγχου και επανεκτέλεσης του αλγορίθμου για τον περιορισμό του λάθους ο αλγόριθμος Gunnar Farneback αποδεικνύεται κατά τους συγγραφείς πιο γρήγορος αν και η τελική επιλογή εξαρτάται επίσης από παράγοντες όπως η υφή των αντικειμένων και το μέγεθος ανάλυσης του βίντεο.

Μία δεύτερη έρευνα (Radhakrishnan, Sharma, & Murthy, 2017) απέδειξε πως στην περίπτωση μεγάλων μετατοπίσεων των περιοχών φωτεινότητας μεταξύ δύο πλαισίων τα αποτελέσματα της μεθόδου Lucas-Kanade παρουσιάζουν απόκλιση σε σχέση με την πραγματική ταχύτητα μετατόπισης αντικειμένων στον τρισδιάστατο χώρο. Αντίθετα, το σφάλμα υπολογισμού περιορίζεται σημαντικά στην περίπτωση του αλγορίθμου του Gunnar Farneback, κάτι που οφείλεται στον υπολογισμό της μετατόπισης για κάθε εικονοστοιχείο εντός των γειτονικών πλαισίων. Ένα δεύτερο συμπέρασμα είναι πως οι δύο αλγόριθμοι οπτικής ροής παράγουν ένα πιο ομαλό πεδίο ταχύτητας σε σχέση με τη μέθοδο Ταχυμετρίας Απεικόνισης Τροχιοδεικτικών Σωματιδίων (Particle Image Velocimetry - PIV) έχοντας ταυτόχρονα μειωμένες υπολογιστικές απαιτήσεις. Με βάση τα αποτελέσματα των (Radhakrishnan, Sharma, & Murthy, 2017, p. 4) ο αλγόριθμος του Gunnar Farneback αξιολογείται ως μία πολύ καλή επιλογή για τη λήψη μετρήσεων αυξημένης ακρίβειας σε σχέση με τον αλγόριθμο Lucas-Kanade και τη χρήση του σε ένα σύστημα του IoT με χαμηλές προδιαγραφές επεξεργασίας.

2.7 Σημαντικότητα της έρευνας

Οι σχετικές έρευνες που έχουν ασχοληθεί με την απόδειξη πλαστογραφίας βίντεο μέσω ανάλυσης της οπτικής ροής έχουν ως επί το πλείστον βασιστεί στη μέθοδο Lucas-Kanade. Τα πλεονεκτήματά τους είναι η αξιοπιστία των αποτελεσμάτων και η υψηλή εφαρμοστικότητα τους καθώς χαρακτηρίζονται από γρήγορη ταχύτητα εκτέλεσης και παρουσιάζουν σχετική ανθεκτικότητα σε συμπίεση του παραποιημένου βίντεο ή σε εισαγωγή θορύβου και παραμορφώσεων από το μέσο λήψης. Η παρούσα έρευνα συμπληρώνει την υπάρχουσα βιβλιογραφία καθώς δεν υπάρχει κατά τη γνώση μας κάποια μέθοδος εύρεσης πλαστογραφίας που να εφαρμόζει τον πιο πρόσφατο αλγόριθμο του Gunnar Farnebäck. Επιπροσθέτως, σύμφωνα με προαναφερθείσα έρευνα, ο αλγόριθμος μπορεί υπό συνθήκες να επιτύχει καλύτερες ταχύτητες εκτέλεσης από τον Lucas-Kanade και επομένως κρίθηκε σημαντικό να διερευνηθεί η αποτελεσματικότητά του. Η μέθοδος που προτείνεται είναι ικανή να ανιχνεύσει τρεις τρόπους χειραγώγησης βίντεο και όπως αποδεικνύεται στο 4^ο κεφάλαιο, είναι ανεξάρτητη από τον τρόπο συμπίεσης του βίντεο. Ένα νέο δεδομένο που εισάγεται είναι υλοποίησή της σε συσκευή IoT (Raspberry Pi 4B) η οποία είναι φθηνή σε κόστος και κατώτερη σε επιδόσεις από τα συστήματα που έχουν χρησιμοποιηθεί για τις παραπάνω μεθόδους. Η αύξηση του διαδικτύου των πραγμάτων καθιστά κάτι τέτοιο ιδιαίτερα σημαντικό ως προς τις δυνατότητες που παρέχονται. Συγχρόνως, η μέθοδος είναι υλοποιήσιμη σε κάθε λειτουργικό σύστημα και λόγω της χρήσης ανοιχτού κώδικα, η ενσωμάτωσή της είναι εφικτή για μελλοντικές έρευνες και παράλληλη χρήση σε εφαρμογές του IoT (Sklavos, Hübner, Göhringer, & Kitsos, 2013). Τέλος, η παρούσα έρευνα μπορεί να αποτελέσει ένα μέσο σύγκρισης των δύο γνωστών αλγορίθμων οπτικής ροής ως προς την εφαρμογή τους στη δικανική πολυμέσων βίντεο.

Κεφάλαιο 3

Σχεδιασμός και Υλοποίηση

Συστήματος

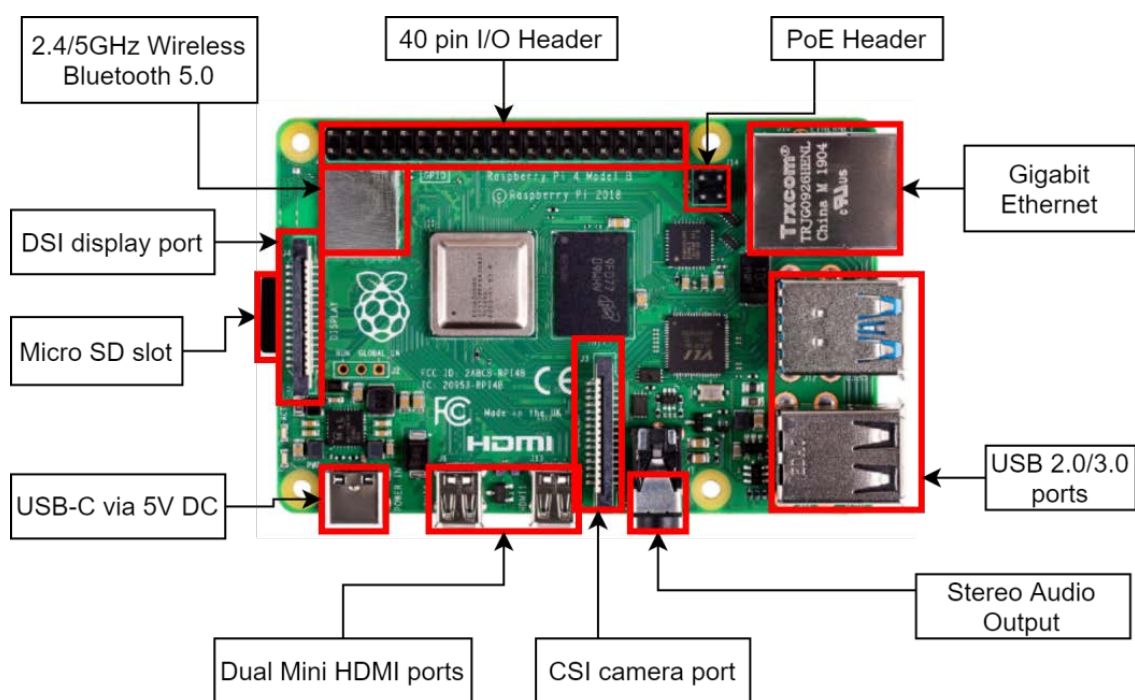
Έχοντας μελετήσει τη σχετική βιβλιογραφία, είναι πλέον εφικτός ο καθορισμός της βάσης για το εργαλείο της μεταπτυχιακής διατριβής. Στο ακόλουθο κεφάλαιο αναλύεται η βασική ιδέα, οι προκλήσεις, οι τεχνολογίες που χρησιμοποιήθηκαν και ο τρόπος υλοποίησης.

3.1 Βασικά δομικά στοιχεία του εργαλείου

3.1.1 Raspberry Pi 4b

Η συσκευή που χρησιμοποιήθηκε ως βάση του συστήματος είναι το Raspberry Pi 4B. Το συγκεκριμένο μοντέλο χρησιμοποιεί τον τετραπύρηνο επεξεργαστή Broadcom BCM2711 αρχιτεκτονικής ARM Cortex-A78 και χρονισμένο στα 1,5GHz (36% ταχύτερος από το μοντέλο 3B+). Το Raspberry Pi 4B χρησιμοποιεί 4GB LPDDR4 μνήμης RAM παρέχοντας σημαντικό χώρο προσωρινής αποθήκευσης σε σχέση με τον προκάτοχό του 3B+ (1GB LPDDR2). Το λειτουργικό σύστημα που εγκαταστάθηκε είναι το Raspberry Pi

OS 32bit (πρώην Raspbian). Είναι σημαντικό να αναφερθεί πως το συγκεκριμένο λειτουργικό σύστημα δεν επιτρέπει στην προκαθορισμένη εγκατάστασή του τον υπερχρονισμό του επεξεργαστή. Αν και υπάρχει ανεπίσημη μέθοδος ώστε να επιτευχθεί ο υπερχρονισμός, κάτι τέτοιο δεν κρίθηκε απαραίτητο καθώς μπορεί να οδηγήσει σε αστάθεια του συστήματος. Επιπλέον, όπως αποδεικνύεται στο Κεφάλαιο 4, η ταχύτητα των 1.5GHz είναι επαρκής για την επεξεργασία αρχείων βίντεο όταν γίνεται βελτιστοποίηση μέσω λογισμικού. Ο αρχικός σχεδιασμός του συστήματος περιλαμβάνει την εγκατάσταση οθόνης 7" και βιντεοκάμερας. Οι συγκεκριμένες προσθήκες εφαρμόστηκαν πειραματικά για την δημιουργία μίας πλήρους αυτόνομης και φορητής συσκευής και είναι προαιρετικές καθώς το τελικό σύστημα επιτρέπει την εκτέλεση του προγράμματος απομακρυσμένα μέσω VNC, SSH και FTP servers. Οι διασυνδέσεις που παρέχονται από το Raspberry Pi 4B για υλικό και δικτύωση φαίνονται στην Εικόνα 3.

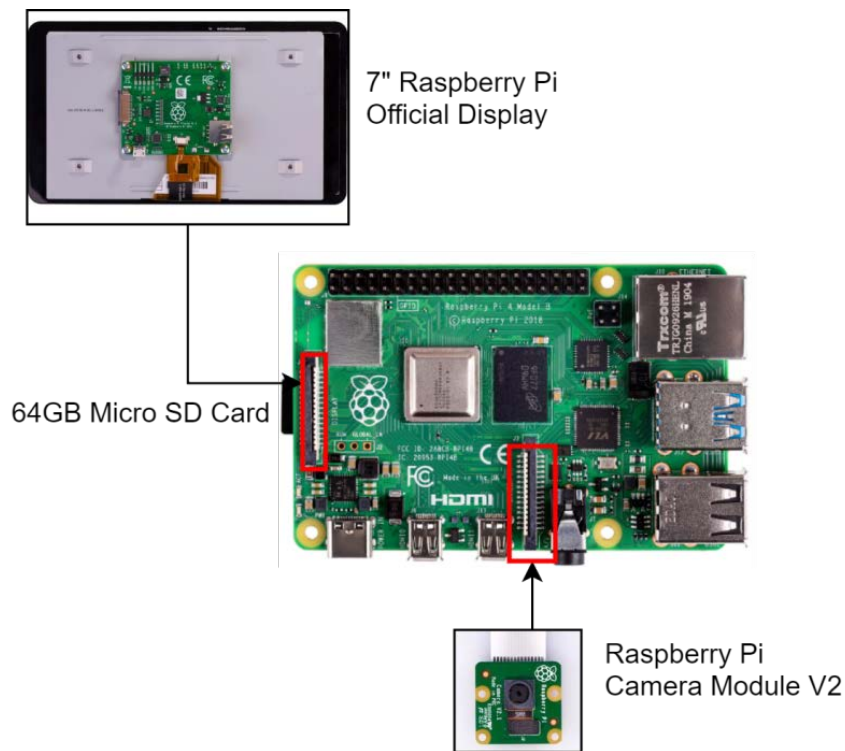


Εικόνα 3. Συνδεσιμότητα στο Raspberry Pi 4B.

Πηγή εικόνας υλικού <https://www.raspberrypi.org>

Η τελική υλοποίηση σε επίπεδο υλικού για την παρούσα μεταπτυχιακή διατριβή περιλαμβάνει μία κάμερα και μία οθόνη 7" για την καλύτερη εποπτεία του γραφικού περιβάλλοντος. Η οθόνη σε πραγματικά σενάρια δεν είναι απαραίτητη καθώς το εργαλείο είναι εκτελέσιμο μέσω VNC server. Η χρήση της στα πλαίσια της έρευνας κρίθηκε σημαντική καθώς η απομακρυσμένη σύνδεση με VNC server δύναται να

οδηγήσει σε πτώση των καρτέ ανά δευτερόλεπτο (fps drops) και δε θα καθιστούσε εφικτή την πλήρη αξιολόγηση του οπτικού αποτελέσματος της επεξεργασίας. Η Εικόνα 4 παρουσιάζει τις συνδέσεις του εξωτερικού υλικού όπως εφαρμόστηκαν στα πλαίσια της μεταπτυχιακής διατριβής.



Εικόνα 4. Σύνδεση κάμερας και οθόνης στο Raspberry Pi 4B της μεταπτυχιακής διατριβής.

Πηγή εικόνων υλικού <https://www.raspberrypi.org>

3.1.2 OpenCV και Python

Στη συγκεκριμένη υποενότητα αναλύονται οι λόγοι που οδήγησαν στη χρήση της OpenCV μέσω της γλώσσας σεναρίων python. Η OpenCV (Open Computer Vision) είναι μία βιβλιοθήκη αλγορίθμων προορισμένων για την ανάλυση βίντεο και εικόνας. Δημιουργήθηκε αρχικά από την Intel και πλέον ο κώδικάς είναι ανοιχτός για τη χρήση εμπορικών και μη εφαρμογών. Η OpenCV αποτελεί ένα API της C++ και μπορεί να αξιοποιηθεί με της χρήση των γλωσσών προγραμματισμού C++, Python, Java και Matlab. Κάτι τέτοιο την καθιστά ιδανική για το εργαλείο της διπλωματικής διατριβής καθώς είναι εφικτή η χρήση ευέλικτου και συμπαγούς κώδικα της Python χωρίς σημαντικούς περιορισμούς στην ταχύτητα εκτέλεσης που ειδήλλως προκύπτουν από τη χρήση μίας γλώσσας σεναρίων. Επιπλέον πλεονέκτημα της βιβλιοθήκης είναι ότι υποστηρίζεται από τα πιο διαδεδομένα λειτουργικά συστήματα (Linux, Mac OS, Windows, Android) και

εμπλουτίζεται συνεχώς με αλγορίθμους τόσο από τη μη κερδοσκοπική οργάνωση OpenCV.org όσο και από ανεξάρτητους προγραμματιστές (opencv-contrib). Η έκδοσή της κατά τη συγγραφή της παρούσας μεταπτυχιακής διατριβής είναι η 4.5.0. Ένα πλεονέκτημα της κλήσης της OpenCV σε κώδικα Python είναι η επιστροφή των συναρτήσεων της σε μορφή πινάκων της βιβλιοθήκης numpy. Οι πίνακες της numpy παρέχουν σημαντική ευελιξία στην επεξεργασία των δεδομένων καθώς υποστηρίζονται από τη πλειοψηφία των βιβλιοθηκών της python που προορίζονται για μαθηματικές πράξεις, μετασχηματισμούς και στατιστική ανάλυση. Όπως παρουσιάζεται στην επόμενη υποενότητα, τα αποτελέσματα του εργαλείου βασίζονται στη χρήση δεδομένων από πίνακα της numpy μέσω της βιβλιοθήκης matplotlib για την οπτικοποίησή τους.

3.1.3 Matplotlib

Η matplotlib αποτέλεσε την επιλογή για τη στατιστική απεικόνιση των τιμών της οπτικής ροής. Είναι σχεδιασμένη για να χρησιμοποιεί πίνακες της numpy και παρέχει την απαιτούμενη ευελιξία ώστε να δημιουργηθούν κατανοητά γραφήματα για την επισκόπηση των αποτελεσμάτων. Η matplotlib είναι βιβλιοθήκη ανοιχτού κώδικα και μπορεί να αντικαταστήσει επαρκώς το περιβάλλον οπτικοποιήσεων της ευρέως χρησιμοποιημένης MATLAB.

3.2 Υλοποίηση

Όπως αναφέρθηκε, το εργαλείο της μεταπτυχιακής διατριβής έχει υλοποιηθεί με τη γλώσσα Python. Η εκτέλεσή του γίνεται σε περιβάλλον γραμμής εντολών χρησιμοποιώντας την εντολή

```
python cmf_detect_ouc.py -run '<path>' -p <float> -l <int> -w <int> -i <int> -n <int> -s <float> -scale <int>
```

Οι παράμετροι που χρησιμοποιούνται κατά την εκτέλεση είναι:

- **run:** Η παράμετρος χρησιμοποιείται για την έναρξη του αλγορίθμου. Αν αντί για -run χρησιμοποιηθεί -h, η έξοδος του προγράμματος είναι μία λίστα πληροφοριών και οδηγιών.
- **path:** Η διαδρομή του βίντεο προς ανάλυση στο μέσο αποθήκευσης.

- **p (pyramid):** Ο δείκτης αποκλιμάκωσης του κάθε πλαισίου για την πυραμίδα Gauss. Προτείνεται η τιμή 0.5 ως βέλτιστη. Για την αύξηση της ταχύτητας εκτέλεσης προτείνονται μικρότερες τιμές ενώ μεγαλύτερες τιμές παρέχουν πιο συμπαγή απεικόνιση της οπτικής ροής.
- **l (levels):** Ο αριθμός επιπέδων της πυραμίδας Gauss. Οι προτεινόμενες τιμές είναι έως 5 για μέγιστη στιβαρότητα ενώ μικρότερες τιμές παρέχουν καλύτερη ταχύτητα εκτέλεσης. Το ελάχιστο που μπορεί να εφαρμοστεί είναι η τιμή 1 όπου δε δημιουργείται κανένα επίπεδο.
- **w (window):** Το μέγεθος του πίνακα NxN για το φίλτρο που παράγεται από τη Γκαουσιανή κατανομή. Μικρότερες τιμές παρέχουν καλύτερο οπτικό αποτέλεσμα με αυξημένο θόρυβο ενώ μεγαλύτερες τιμές παράγουν ένα θολό αποτέλεσμα αλλά πιο συμπαγές και με μικρότερο ποσοστό θορύβου.
- **i (iterations):** Ο αριθμός επαναλήψεων της πολυωνυμικής επέκτασης για κάθε επίπεδο της πυραμίδας Gauss. Μικρότερες τιμές επιτρέπουν ταχύτερη εκτέλεση του αλγορίθμου αλλά μπορούν να οδηγήσουν σε σφάλμα υπολογισμού των γειτονικών εικονοστοιχείων. Μεγαλύτερες τιμές περιορίζουν το σφάλμα και προσδίδουν ακρίβεια στον υπολογισμό της οπτικής ροής. Η προτεινόμενη μέση τιμή είναι 3.
- **n (polynomial number):** Ο βαθμός του πολυωνύμου για τον υπολογισμό επέκτασης στα γειτονικά εικονοστοιχεία. Οι τυπικές τιμές είναι 5 για μεγαλύτερη ταχύτητα και 7 για καλύτερη πυκνότητα.
- **s (polynomial sigma):** Η τυπική απόκλιση Gauss. Προτείνεται η χρήση της τιμής 1.1 για $n=5$ ή 1.5 για $n \in [6,7]$.
- **scale:** Δεν αποτελεί παράμετρο του αλγορίθμου οπτικής ροής του Gunnar Farneback αλλά χρησιμοποιείται για την αρχική αποκλιμάκωση των 2 πλαισίων που εισάγονται στον αλγόριθμο. Μικρότερες τιμές αυξάνουν την ταχύτητα εκτέλεσης αλλά είναι απαραίτητο να υπάρχει συνοχή μεταξύ της ανάλυσης των εισαγόμενων πλαισίων και των παραμέτρων του αλγορίθμου.

Αρχικά εκτελούνται οι έλεγχοι για την ορθότητα των παραμέτρων και την ύπαρξη του βίντεο. Είναι σημαντικό να ελέγχονται πάντα τα πλαίσια εισόδου ώστε να αποφευχθεί κάποιο σφάλμα στην εκτέλεση. Σε κάθε επανάληψη του βασικού κόμβου του εργαλείου εκτελούνται τα ακόλουθα βήματα.

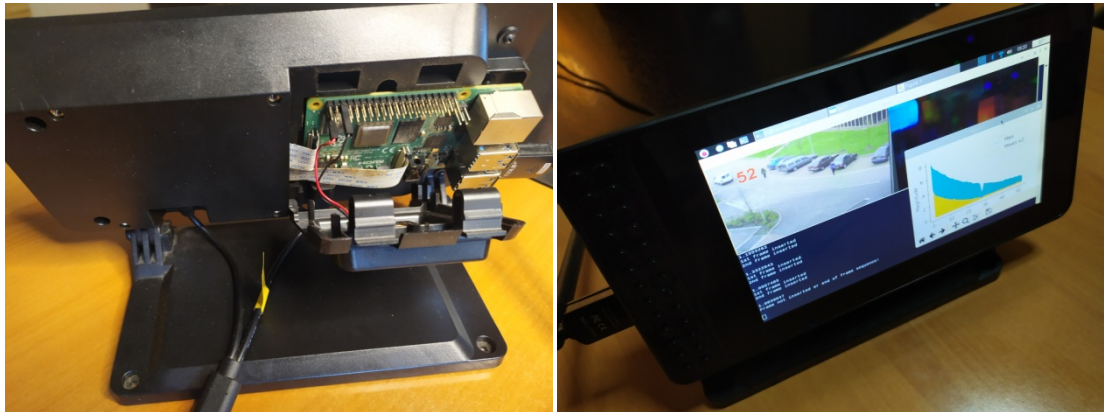
- 1) Εισάγονται τα δύο πλαίσια σε μεταβλητές.

- 2) Εφαρμόζονται μετασχηματισμοί αποκλιμάκωσης (downscale) και αποχρωματισμού των καναλιών RGB σε κλίμακα του γκρι. Στόχος είναι να μειωθούν οι υπολογιστικές απαιτήσεις περιορίζοντας τον αριθμό των εικονοστοιχείων ανά πλαίσιο και μειώνοντας τον αριθμό των τριών καναλιών χρώματος σε ένα.
- 3) Ορίζεται ένας πίνακας τύπου HSV (Hue, Saturation, Value) με μέγεθος ίδιο με αυτό των πλαισίων που εισήχθησαν για τη χρωματική αποτύπωση του μεγέθους και γωνίας της οπτικής ροής. Ο λόγος που χρησιμοποιείται το μοντέλο HSV και όχι RGB είναι γιατί αποδίδει τη χρωματική τιμή βάσει της κλίσης που έχει ένα αντικείμενο. Συγκεκριμένα οι χρωματικές τιμές σε σχέση με την κλίση ενός αντικειμένου αναφέρονται στον Πίνακα 1.
- 4) Κλήση της μεθόδου `cv2.calcOpticalFlowFarneback()` για τον υπολογισμό της οπτικής ροής στα δύο πλαίσια. Η μέθοδος επιστρέφει δύο πίνακες με το μέγεθος (magnitude) και την κλίση (angle) για κάθε εικονοστοιχείο εντός των δύο πλαισίων.
- 5) Ενημερώνονται οι δύο λίστες `magnitude_max` και `magnitude_mean` με τα μέσα και μέγιστα μεγέθη της οπτικής ροής.
- 6) Εφαρμόζονται μετασχηματισμοί ώστε να αποδοθούν οι σωστές τιμές σε κάθε θέση του πίνακα HSV και στη συνέχεια ο πίνακας μετασχηματίζεται σε πίνακα RGB για την προβολή των πλαισίων.

Απόχρωση	Ελάχιστη γωνία κλίσης	Μέγιστη γωνία κλίσης
Κόκκινο	0°	60°
Κίτρινο	61°	120°
Πράσινο	121°	181°
Κυανό	181°	240°
Μπλε	241°	300°
Πορφυρό	301°	360°

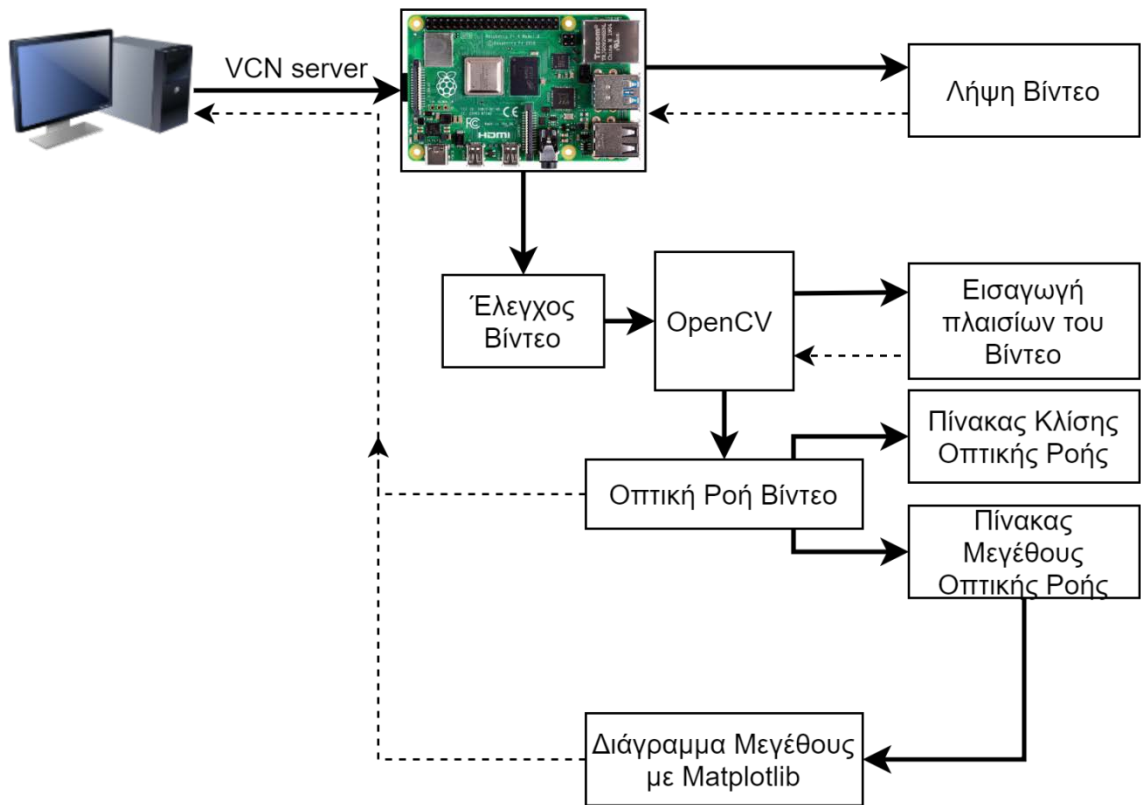
Πίνακας 1. Πίνακας χρωματικών τιμών σε αντιστοιχία με τη γωνία κλίσης στο μοντέλο HSV.

Με τη λήξη της ανάλυσης και προβολής χρησιμοποιείται η matplotlib για την παραγωγή του γραφήματος αντιστοίχισης πλαισίων σε μέγιστες και μέσες τιμές των μεγεθών της οπτικής ροής ανά ζεύγος πλαισίων. Τέλος, προβάλλεται το παραγόμενο σχεδιάγραμμα όπου παρέχεται η δυνατότητα μεγέθυνσης περιοχών για πιο ακριβή ανάλυση. Το πρόγραμμα αναμένει από το χρήστη να πατήσει το πλήκτρο “q” για να εξέλθει. Στην Εικόνα 5 φαίνεται η τελική υλοποίηση του εργαλείου κατά την εκτέλεση στο Raspberry Pi 4B.



Εικόνα 5. Υλοποίηση του συστήματος και εκτέλεση του εργαλείου σε Raspberry Pi 4B

Στο Σχήμα 3 παρουσιάζεται συνοπτικά η λειτουργία του εργαλείου. Τα διακεκομμένα βέλη αφορούν την έξοδο της επεξεργασίας. Ο χρήστης συνδέεται μέσω ενός VNC client (στο πλαίσιο της μεταπτυχιακής διατριβής χρησιμοποιήθηκε το RealVNC). Προαιρετικά το Raspberry Pi μπορεί να πραγματοποιήσει λήψη βίντεο και αποθήκευση στη Micro SD κάρτα. Ο έλεγχος του βίντεο εισάγει τα πλαίσια του βίντεο ανά δύο στην OpenCV και υπολογίζεται η οπτική ροή. Ο αλγόριθμος επιστρέφει ως έξοδο το οπτικό αποτέλεσμα της οπτικής ροής ενημερώνοντας παράλληλα δύο πίνακες της numpy με το μέγεθος και την κλίση ανά εικονοστοιχείο. Σε κάθε κύκλο επεξεργασίας η μέγιστη και μέση τιμή του πίνακα μεγεθών εισάγονται σε δύο διαφορετικές λίστες (magnitude_max, magnitude_mean) οι οποίες χρησιμοποιούνται από τη Matplotlib για την παραγωγή του γραφήματός μεγεθών της οπτικής ροής. Με τη λήξη της επεξεργασίας το διάγραμμα μεγεθών παρουσιάζεται στο χρήστη για την ανάλυση και ανίχνευση ομοιοτήτων ή απότομων αυξομειώσεων των μεγεθών.



Σχήμα 3. Σχηματική αναπαράσταση της λειτουργίας του εργαλείου

Κεφάλαιο 4

Εκτέλεση & Αποτελέσματα

Το επόμενο βήμα μετά από το σχεδιασμό και την υλοποίηση του εργαλείου της μεταπτυχιακής διατριβής είναι η εκτέλεσή του σε περιβάλλον σεναρίων. Το παρών κεφάλαιο περιλαμβάνει την παρουσίαση του εργαλείου μέσα από έναν αριθμό πειραμάτων με σκοπό να αξιολογηθεί η αποτελεσματικότητά του. Επιπλέον μέσω της εκτέλεσής του είναι εφικτή η σύγκριση των αποτελεσμάτων με την αρχική ιδέα ώστε να επιβεβαιωθεί το κατά πόσο έχει επιτευχθεί ο στόχος της έρευνας.

4.1 Παράμετροι και σύνολα βίντεο

Τα ακόλουθα πειράματα πραγματοποιήθηκαν με τη χρήση έτοιμων συνόλων από πλαστογραφημένα βίντεο. Τα βίντεο που ανήκουν στο REWIND project (Bestagini, Milani, Tagliasacchi, & Tubaro, 2013), (Qadir, Yahaya, & T.S Ho, 2012) χρησιμοποιήθηκαν για την εύρεση αντιγραφής και ανακατανομής πλαισίων σε χωρικό και χρονικό επίπεδο. Περιλαμβάνονται επομένως πλαστογραφίες εντός ομάδων πλαισίων (copy-move) και περιοχών από ομάδες πλαισίων (splicing). Τα βίντεο του CVIP Group (Ardizzone & Mazzola, 2015) έχουν επίσης υποστεί αλλαγές σε χωρικό και χρονικό επίπεδο αλλά εμφανείς στο ανθρώπινο μάτι σε αντίθεση με αυτά του REWIND

project τα οποία πληρούν προδιαγραφές ιδανικών συνθηκών πλαστογραφίας. Με γνώμονα ότι η πλαστογραφία βίντεο συχνά πραγματοποιείται από μη ειδικούς της επεξεργασίας πολυμέσων (multimedia professionals), θεωρήθηκε πως τα βίντεο του CVIP group μπορούν προσομοιώσουν πραγματικές περιπτώσεις αλλοίωσης και επομένως έγινε χρήση τους.

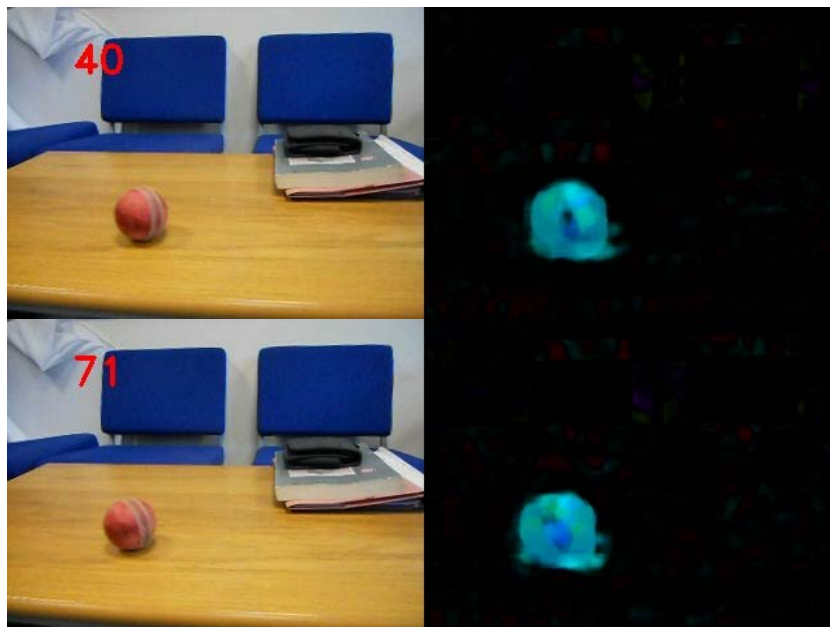
Για τα δύο σύνολα χρησιμοποιήθηκαν διαφορετικές παράμετροι στον αλγόριθμο του Farnebäck. Συγκεκριμένα, για το σύνολο βίντεο του REWIND project ο αλγόριθμος χρησιμοποιεί 5 επίπεδα πυραμίδας Gauss με 50% αποκλιμάκωση σε κάθε επίπεδο. Επιπλέον, για κάθε επίπεδο ο αλγόριθμος εκτελείται 3 φορές για τη μείωση του σφάλματος και οι δείκτες `poly_z` και `poly_sigma` έχουν τιμές 7 και 1.5 αντίστοιχα. Οι συγκεκριμένες παράμετροι παρέχουν μεγάλη στιβαρότητα και ακρίβεια στον υπολογισμό της οπτικής ροής περιορίζοντας στο ελάχιστο δυνατό την καθυστέρηση εκτέλεσης του αλγορίθμου. Παρατηρήθηκε πως αυξάνοντας περεταίρω τα επίπεδα της πυραμίδας ή τις επαναλήψεις ανά επίπεδο τα αποτελέσματα παρουσιάζουν ελάχιστη βελτίωση σε αναλογία με την αυξημένη υπολογιστική βαρύτητα. Η τιμή του `window` είναι 50 η οποία πρακτικά είναι πολύ μεγάλη για την απόδοση ενός βέλτιστου οπτικού αποτελέσματος. Βασικό πλεονέκτημα της χρήσης μεγάλων τιμών παραθύρου είναι πως αυξάνοντας κατά πολύ το μέγεθός του περιορίζεται ο θόρυβος της οπτικής ροής και το αποτέλεσμα είναι πιο στιβαρό για στατιστική ανάλυση. Το πείραμα 4.2.1 περιλαμβάνει μία σύγκριση των αποτελεσμάτων με μικρό και μεγάλο παράθυρο (τιμές 16 και 50 αντίστοιχα) και επιβεβαιώνει την παραπάνω επιλογή. Για το σύνολο βίντεο του GVIP group περιορίστηκε η στιβαρότητα του αλγορίθμου ώστε να επιτευχθεί ο βέλτιστος χρόνος εκτέλεσης. Ο λόγος είναι πως απότομες αλλαγές στην οπτική ροή οδηγούν σε απότομες και μεγάλες αλλαγές των μέγιστων τιμών μεγέθους της και δεν επηρεάζονται σημαντικά από τον εισαγόμενο θόρυβο ή τη διπλή συμπίεση. Σε περίπτωση που ο αλγόριθμος πρέπει να εφαρμοστεί ώστε να καλύπτει και τις δύο περιπτώσεις ταυτόχρονα, προτείνεται η εκτέλεσή του δύο φορές χρησιμοποιώντας σε κάθε περίπτωση διαφορετικές παραμέτρους για τον έλεγχο και των τριών μορφών πλαστογραφίας (copy-move, delete, splice).

4.2 Πειράματα

Για την αξιολόγηση του εργαλείου πραγματοποιήθηκαν 9 πειράματα σε δύο διαφορετικά σύνολα βίντεο. Όλα τα βίντεο περιλαμβάνουν στατικό φόντο και κίνηση συγκεκριμένων αντικειμένων στο χώρο. Για να οδηγηθούμε σε πιο στιβαρά αποτελέσματα απαιτείται ένας μεγαλύτερος αριθμός πειραμάτων τα οποία θα εισάγουν περισσότερα κινούμενα αντικείμενα.

4.2.1 Λείες μπάλες σε γραφείο

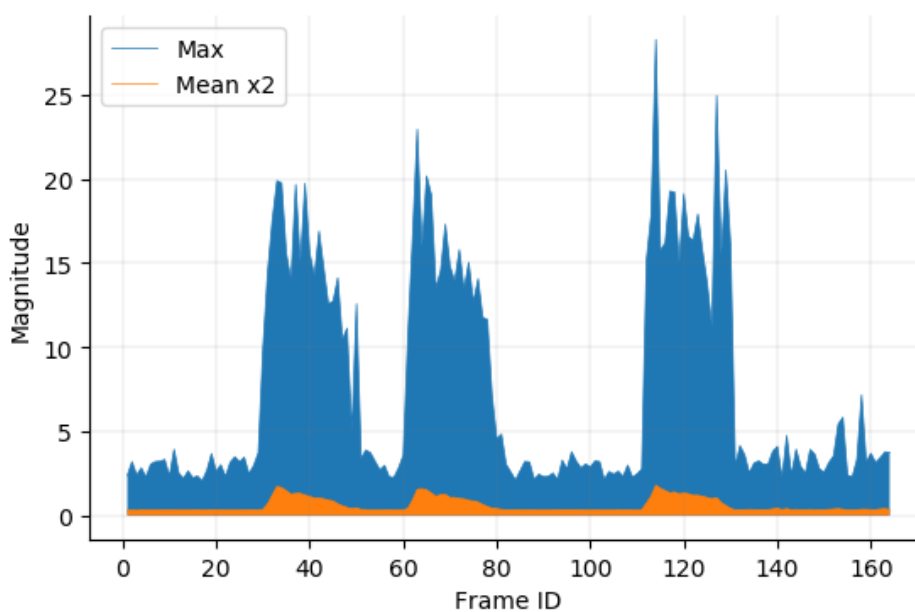
Στο συγκεκριμένο βίντεο (Εικόνα 6) παρατηρούνται 3 μπάλες λείας επιφάνειας σε διαδοχή με οριζόντια αριστερόστροφη κίνηση πάνω σε λεία επιφάνεια γραφείου. Στο πραγματικό βίντεο εμφανίζονται 2 μπάλες και είναι αυτές που φαίνονται στα πλαίσια 40 και 122 της επόμενης εικόνας. Η δεύτερη μπάλα εμφανίζεται μεταξύ των πλαισίων 60-82 μέσω αντιγραφής και επανατοποθέτησης από σημεία περιοχής των πλαισίων 30-47. Το συγκεκριμένο πείραμα αρχικά παρουσιάζει τα αποτελέσματα της ανάλυσης όταν γίνεται χρήση μικρού παραθύρου ($window=16$) για τον αλγόριθμο του Farnebäck.





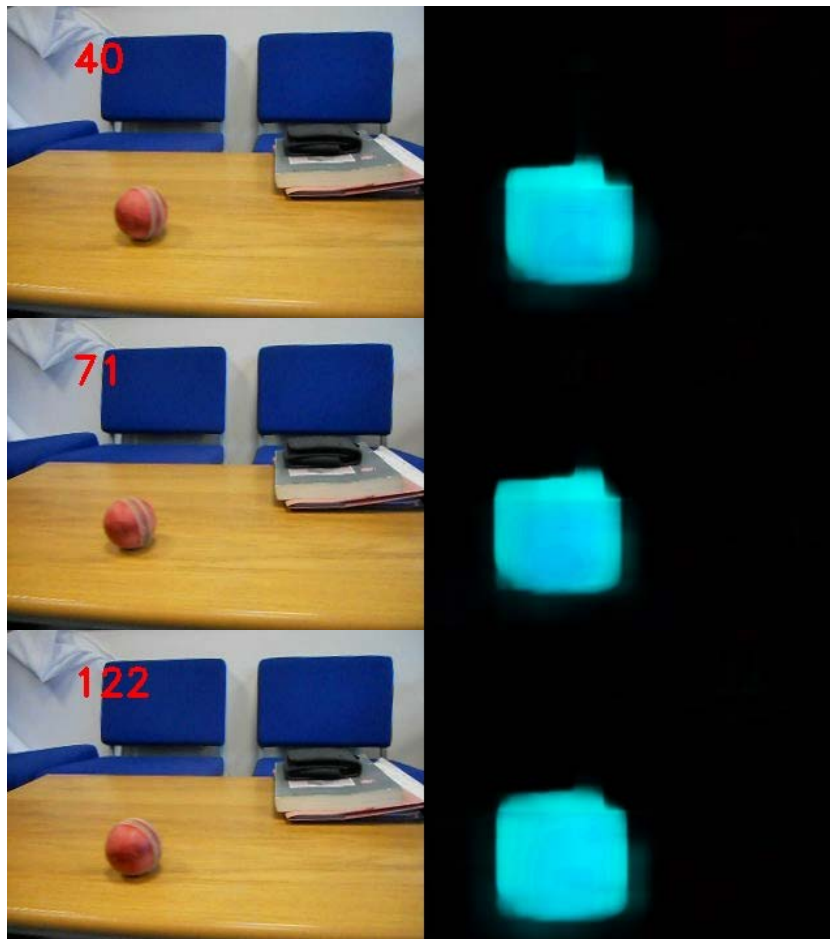
Εικόνα 6. Οπτική ροή με χαμηλή τιμή παραθύρου.

Το οπτικό αποτέλεσμα είναι αντιπροσωπευτικό των αντικειμένων στο βίντεο αλλά οι τιμές της οπτικής ροής δεν είναι επαρκείς για την αναγνώριση πλαστογραφίας όπως φαίνεται στο Διάγραμμα 1. Λόγω του μειωμένου παραθύρου εισάγεται ένα ποσοστό θορύβου από περιοχές με πολύ μικρή κίνηση (θόρυβος κάμερας ή δονήσεις), το οποίο επηρεάζει το συνολικό μέσο και μέγιστο μέγεθος της οπτικής ροής.



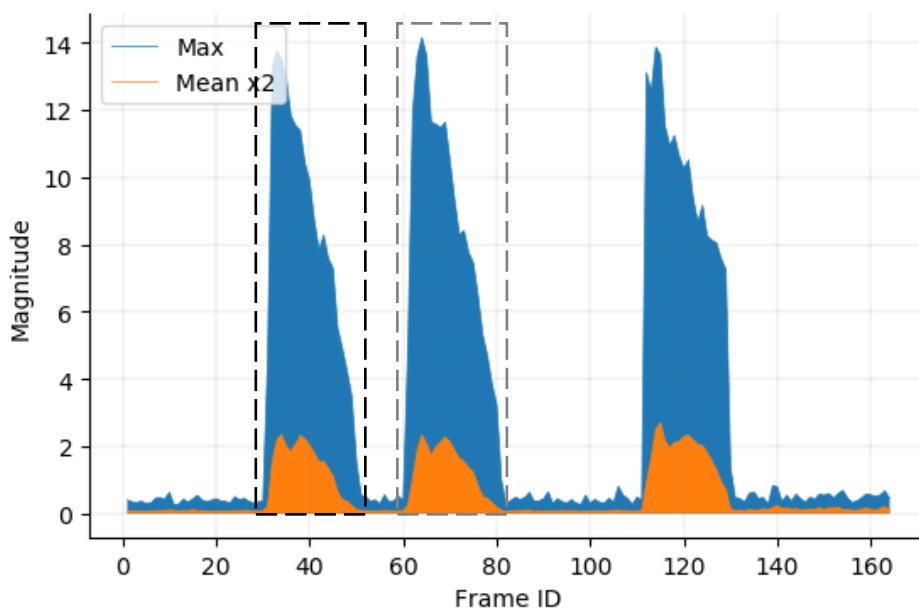
Διάγραμμα 1. Στατιστική προβολή της οπτικής ροής με window=16.

Αυξάνοντας το μέγεθος παραθύρου (window=50) το οπτικό αποτέλεσμα είναι θολό καθώς η οπτική ροή περιλαμβάνει εικονοστοιχεία γύρω από την περιοχή του αντικειμένου που μπορεί να παρουσιάζουν μεταβολές λόγω μικρών αλλαγών φωτεινότητας. Το αποτέλεσμα φαίνεται στην Εικόνα 7.



Εικόνα 7. Οπτική ροή με υψηλή τιμή παραθύρου.

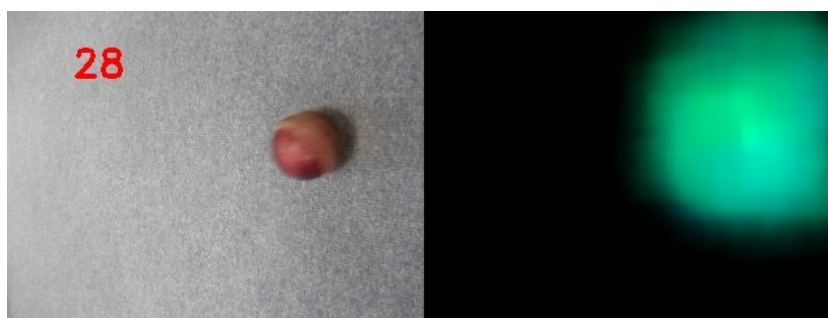
Ένα θετικό αυτής της επιλογής είναι ότι ο επιπλέον θόρυβος της προηγούμενης περίπτωσης έχει εξαφανιστεί και τα δεδομένα που παράγονται έχουν μεγαλύτερο βαθμό συσχέτισης ώστε να αναγνωριστεί η πλαστογραφία όπως φαίνεται στο Διάγραμμα 2. Για την καλύτερη παρουσίαση οι ύποπτες περιοχές έχουν σημειωθεί με περίγραμμα.

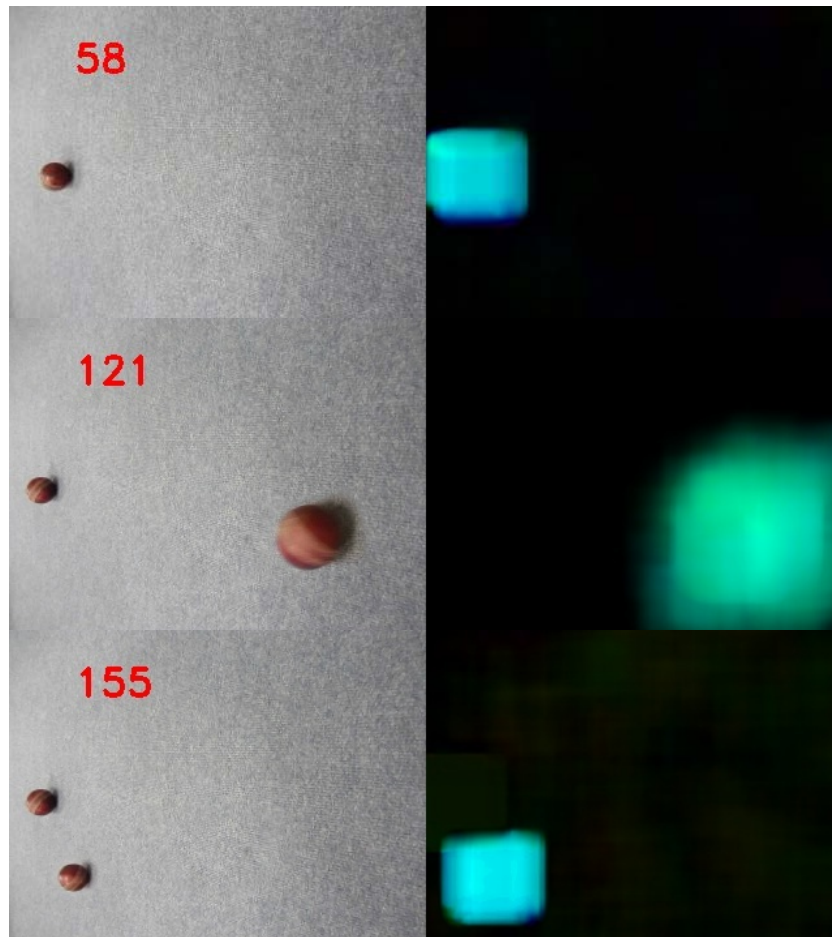


Διάγραμμα 2. Στατιστική προβολή της οπτικής ροής με window=50.

4.2.2 Λείες μπάλες σε τραχύ δάπεδο

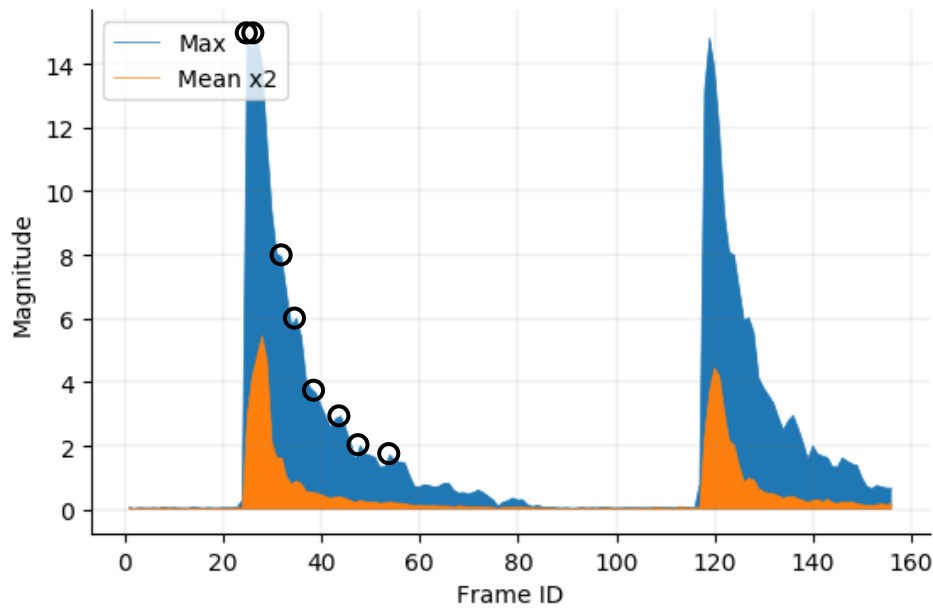
Στο συγκεκριμένο πείραμα εμφανίζονται δύο λείες μπάλες με οριζόντια κίνηση πάνω σε τραχιά επιφάνεια δαπέδου με διαδοχή. Η οπτική ροή επηρεάζεται σε μεγάλο βαθμό από την επιφάνεια των αντικειμένων καθώς υπάρχει διαφοροποίηση στην αντανάκλαση φωτεινών πηγών και στη σκίαση. Σε αντίθεση με το παράδειγμα 4.2.1, η διαφορά φωτεινότητας μεταξύ των κινούμενων αντικειμένων και της σταθερής επιφάνειας είναι πολύ μεγαλύτερη, κάτι που βελτιστοποιεί το αποτέλεσμα του αλγορίθμου αναγνώρισης. Στο πραγματικό βίντεο, μόνο η πρώτη μπάλα εμφανίζεται ενώ η δεύτερη μπάλα προκύπτει με αντιγραφή και επανατοποθέτηση σημείων περιοχών από τα πλαίσια 22-80. Το ανθρώπινο μάτι δε μπορεί να αντιληφθεί τη διαφορά λόγω της αίσθησης του βάθους που κάνει τις μπάλες να φαίνονται διαφορετικές (Εικόνα 8).





Εικόνα 8. Τα πλαίσια 28 και 58 είναι αυθεντικά. Τα πλαίσια 121 και 155 είναι πλαστά.

Όπως φαίνεται στο Διάγραμμα 3 της Εικόνα 8 τα δύο αντικείμενα είναι πραγματικά τα ίδια και χαρακτηρίζονται από τα ίδια δεδομένα οπτικής ροής. Επιπλέον παρουσιάζουν σημαντικές ομοιότητες στις τιμές των κορυφών οπτικής ροής οι οποίες έχουν τις ίδιες αποστάσεις (± 1 πλαίσιο) μεταξύ τους στις δύο διακυμάνσεις ροής. Συγκεκριμένα οι κορυφές που σημειώνονται με κύκλο στη πρώτη διακύμανση (πλαίσια 22-83) έχουν ίσες αποστάσεις με τις αντίστοιχες της δεύτερης (πλαίσια 118-155). Όπως αποδεικνύεται στην ανάλυση των συνολικών αποτελεσμάτων στο τέλος του κεφαλαίου, αυτός ο παράγοντας μπορεί να αποτελέσει τη βάση για το σχεδιασμό αλγορίθμου αυτόματης ανίχνευσης της πλαστογραφίας και του τύπου της.

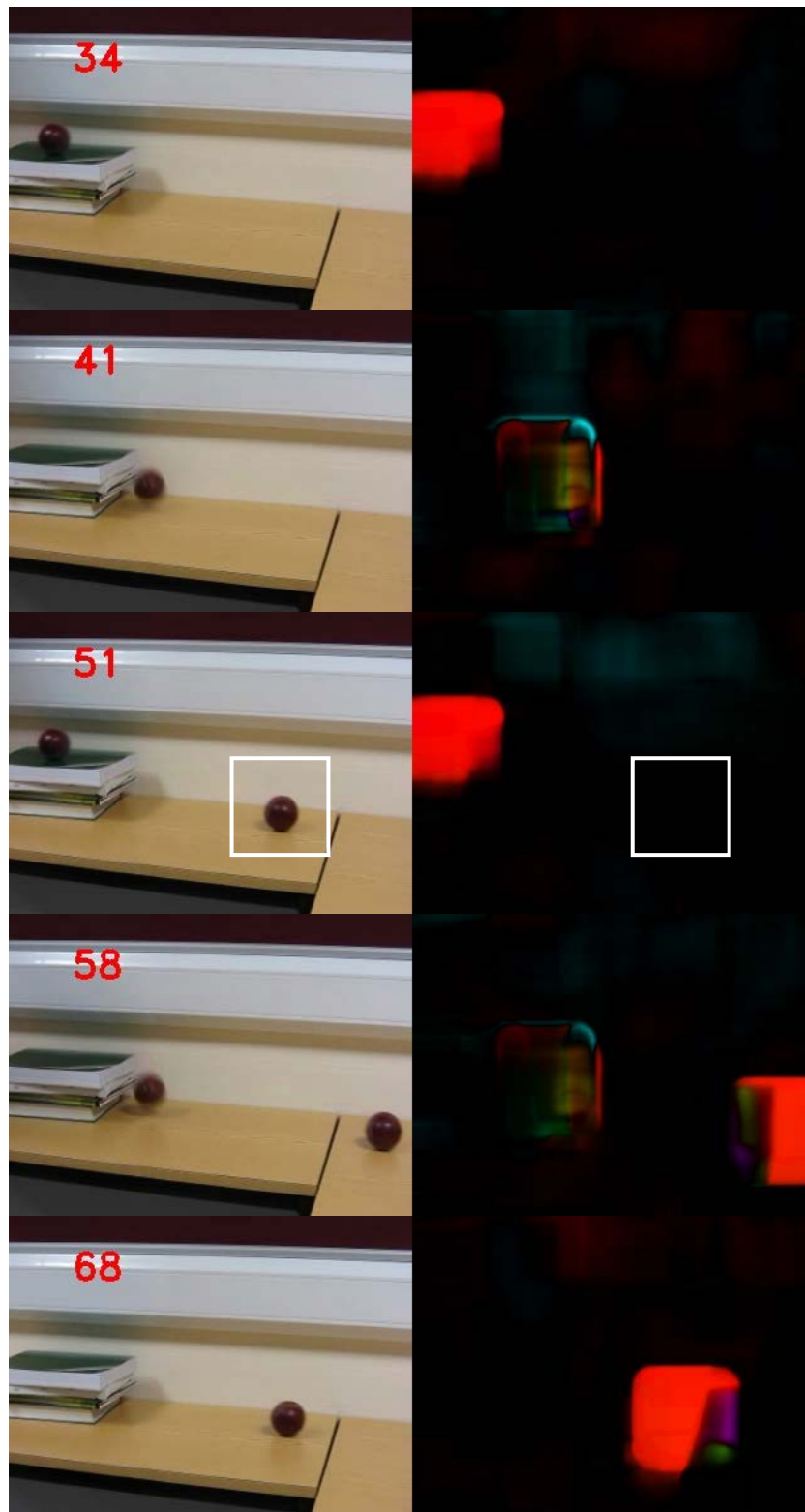


Διάγραμμα 3. Στατιστικές ομοιότητες της οπτικής ροής για τις δύο μπάλες.

4.2.3 Λείες μπάλες σε γραφείο με κίνηση πτώσης

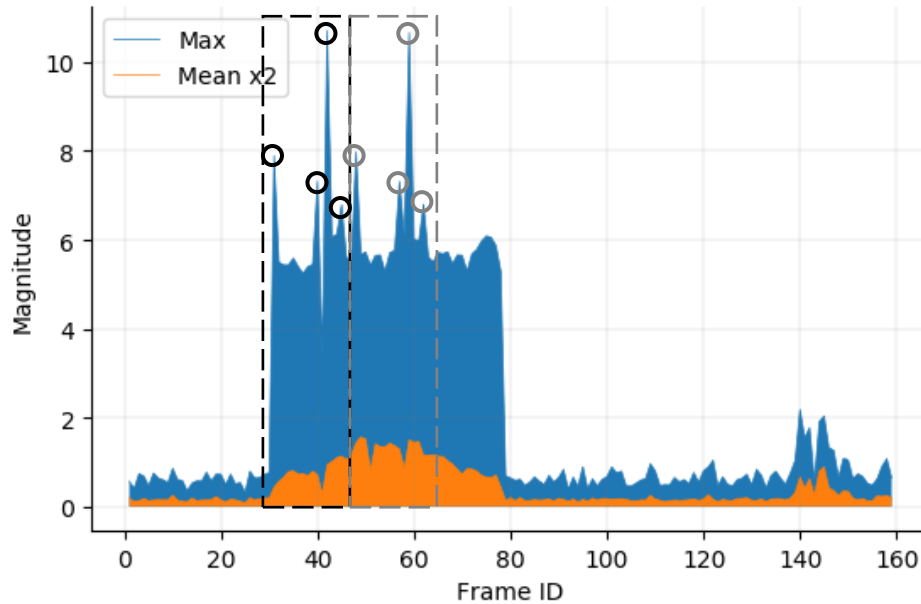
Μία σημαντική διαφορά του συγκεκριμένου πειράματος με τα υπόλοιπα είναι ότι υπάρχουν δύο σύνολα πλαισίων όπου τα αντικείμενα κάνουν απότομη πτώση. Ένα βασικό πρόβλημα των αλγορίθμων οπτικής ροής είναι η ανάγκη για ομαλή μετατόπιση των φωτεινών περιοχών μεταξύ περιοχών εικονοστοιχείων σε γειτονικά πλαίσια. Αν και η πυραμίδα Gauss συχνά επιλύει αυτό το πρόβλημα, πρέπει ανάλογες περιπτώσεις να ληφθούν υπόψη ώστε να αποδειχθεί η αξιοπιστία ενός αλγορίθμου. Ένα εργαστηριακό περιβάλλον επιτρέπει τον έλεγχο των πειραμάτων αλλά κάτι τέτοιο δεν είναι πάντα αντίστοιχο με πραγματικές συνθήκες όπου η ταχύτητα μπορεί να μεταβληθεί απότομα και σημαντικά. Για παράδειγμα, μία σύγκρουση θα οδηγήσει συχνά σε απότομες μεταβολές στην οπτική ροή και την ταχύτητα μετατόπισης αντικειμένων. Στο συγκεκριμένο βίντεο παρουσιάζονται δύο λείες μπάλες με οριζόντια και κάθετη κίνηση οι οποίες εμφανίζονται σε διαδοχή πάνω σε λεία επιφάνεια γραφείου. Η πλαστογραφία γίνεται με αντιγραφή περιοχής των πλαισίων 30-46 και αναδιάταξη. Σημαντικό είναι πως η οπτική ροή για την πρώτη μπάλα χάνεται στο πλαίσιο 51 όπως φαίνεται στην Εικόνα 9. Κάτι τέτοιο είναι συχνά αναμενόμενο στην πραγματικότητα. Επιπλέον παρατηρείται πως ο αλγόριθμος του Farneback είναι στιβαρός σε απότομη αυξομείωση της κίνησης όπως φαίνεται στα πλαίσια 41 και 58. Η οπτική ροή δε χάνεται και ταυτόχρονα ο υπολογισμός της παράγει σωστά

αποτελέσματα με μερική διαφοροποίηση στις τιμές φωτεινότητας, πιθανώς λόγω της επεξεργασίας τους για τη δεύτερη μπάλα.



Εικόνα 9. Αριστερά τα πλαστά πλαίσια. Δεξιά η αντίστοιχη οπτική ροή.

Η πλαστογραφία μπορεί να διακριθεί στο Διάγραμμα 4. Επιπλέον παρατηρείται ομοιότητα το προηγούμενο πείραμα για τις αποστάσεις των κορυφώσεων της οπτικής ροής για τις δύο περιοχές που διαχωρίζονται.

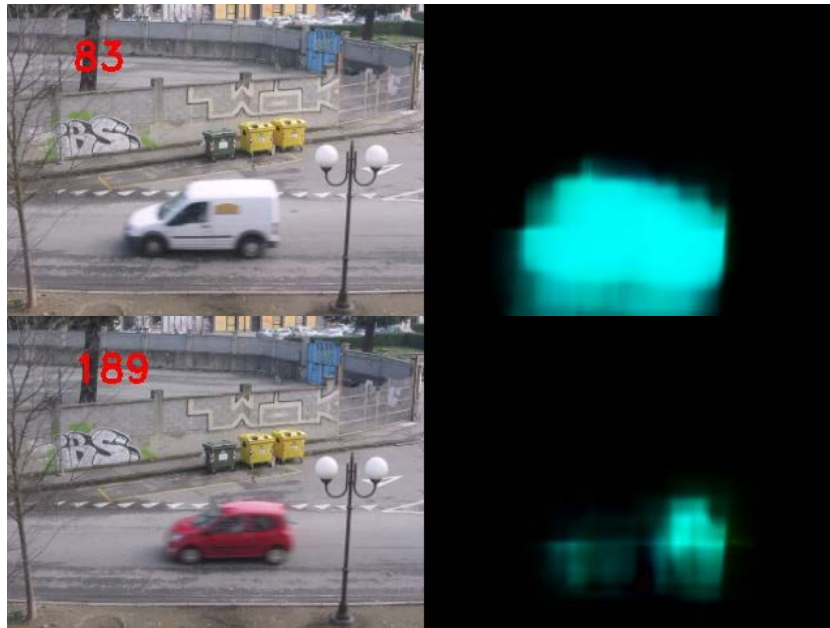


Διάγραμμα 4. Στατιστική προβολή της οπτικής ροής.
Η αρχική παρατήρηση για τις αποστάσεις τους επιβεβαιώνεται.

4.2.4 Αυτοκίνητα διαφορετικού τύπου σε οριζόντια κίνηση

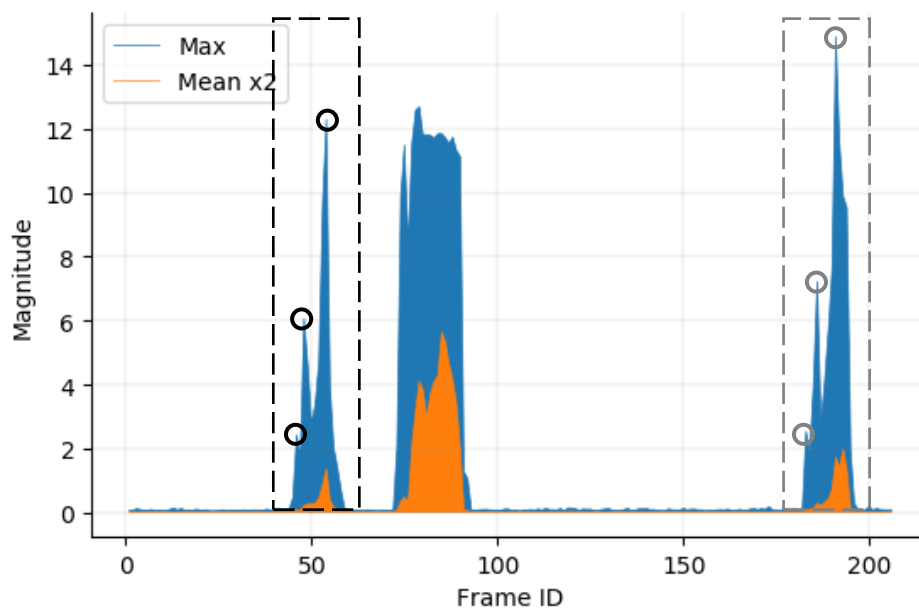
Στο συγκεκριμένο βίντεο παρατηρούνται 3 αυτοκίνητα σε οριζόντια κίνηση με διαδοχή. Στο πραγματικό βίντεο το λευκό Βαν εμφανίζεται πρώτο και ακολουθεί το κόκκινο Ι.Χ. Η πλαστογραφία εφαρμόστηκε σε περιοχές εντός των πλαισίων όπου εμφανίζεται το κόκκινο Ι.Χ με αντιγραφή και αναδιάταξη στην αρχή του βίντεο.





Εικόνα 10. Αριστερά τα πλαίσια του πλαστού βίντεο. Δεξιά η αντίστοιχη οπτική ροή.

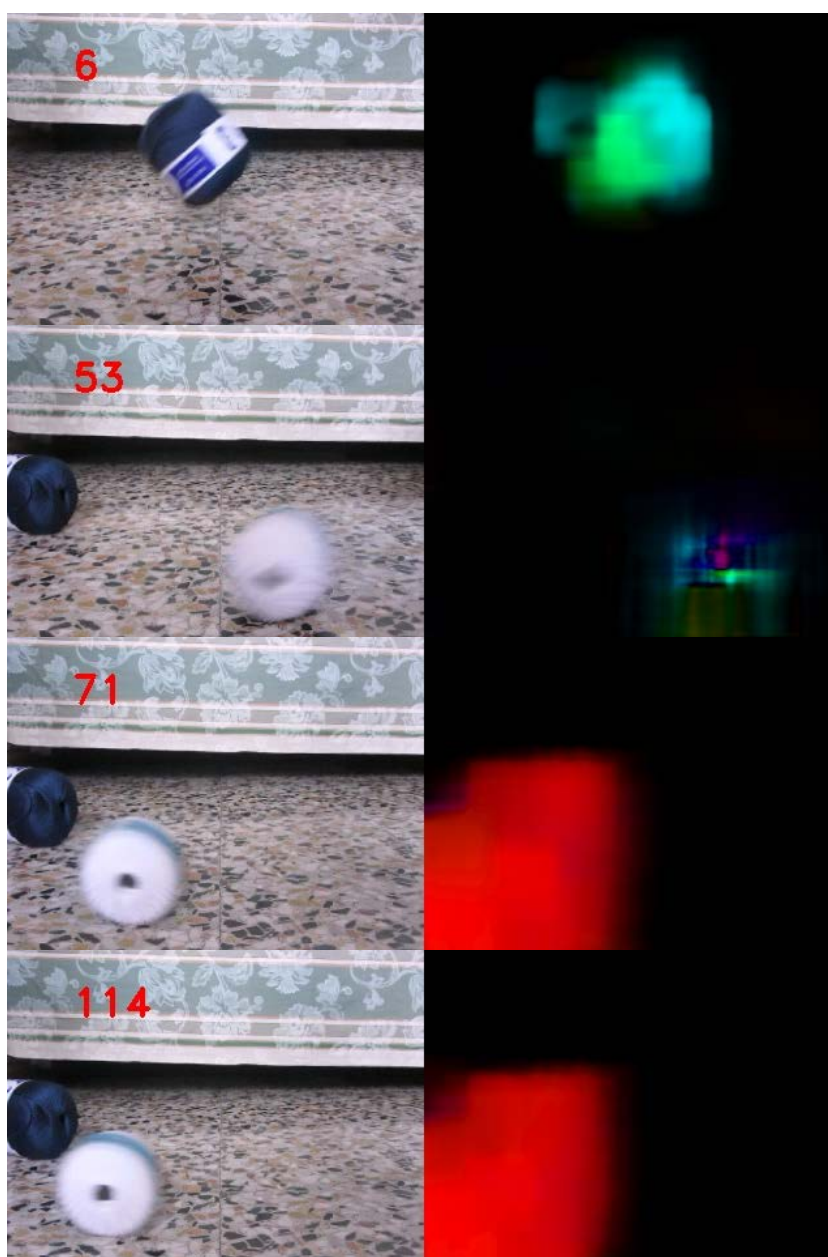
Επιπροσθέτως, παρατηρούνται αλλαγές στις τιμές φωτεινότητας μεταξύ των δύο εμφανίσεων του I.X (Εικόνα 10) κάτι που επηρεάζει το αποτέλεσμα του σχεδιαγράμματος της οπτικής ροής. Αρχικά μπορεί να παρατηρηθεί (Διάγραμμα 5) ότι οι δύο διακυμάνσεις είναι αναλογικά όμοιες αλλά όχι ίσες. Επίσης, σημαντικό είναι πως οι κορυφές της οπτικής ροής για τις δύο εμφανίσεις του I.X έχουν ίσες αποστάσεις μεταξύ τους. Επιβεβαιώνεται ξανά ότι το αποτέλεσμα είναι ικανό να παρέχει μία βάση συσχέτισης για δύο εμφανίσεις του ίδιου αντικειμένου παρά τις διαφορές στη φωτεινότητα και αντίθεση (contrast).



Διάγραμμα 5. Ομοιότητες στη στατιστική προβολή της οπτικής ροής για τα δύο I.X.

4.2.5 Μπάλες κλωστής σε λείο δάπεδο

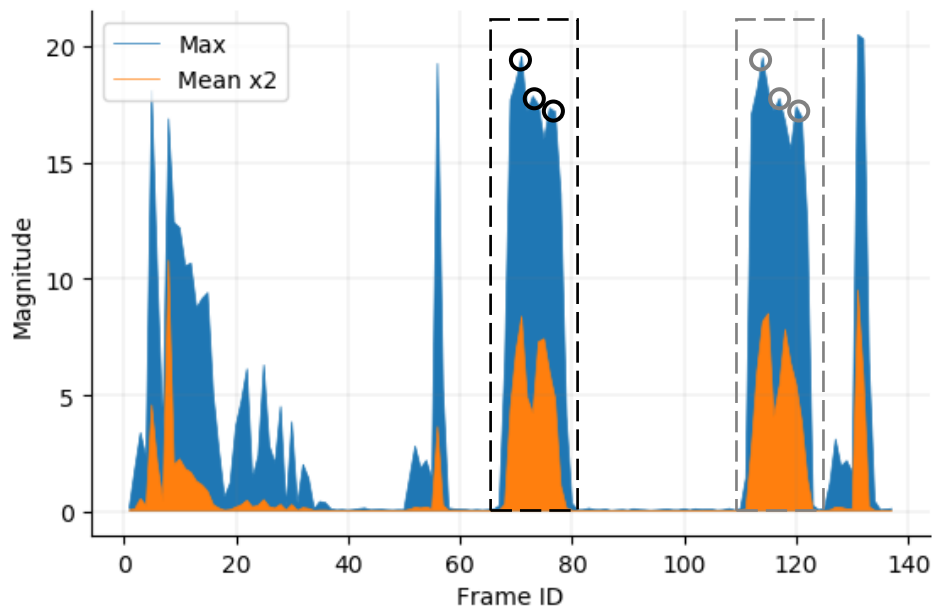
Το ακόλουθο πείραμα κρίθηκε σημαντικό για αρκετούς λόγους. Πρώτον, το πραγματικό βίντεο περιλαμβάνει μπάλες κλωστής που κινούνται απότομα και με μεγάλη ταχύτητα σε λεία επιφάνεια δαπέδου (Εικόνα 11) και επομένως είναι κρίσιμο για να παρατηρηθεί η όποια απώλεια οπτικής ροής. Κατά δεύτερο, τα αντικείμενα έχουν άμεση διαδοχή στην εμφάνισή τους με αποτέλεσμα να είναι δύσκολη η διάκριση της πλαστογραφίας από το ανθρώπινο μάτι. Τέλος, όπως φαίνεται στο Διάγραμμα 6, υπάρχει αυξημένη πιθανότητα παραγωγής false-positive για αντικείμενα των πλαισίων 53 και 130 και των αντίστοιχων ομάδων πλαισίων στις οποίες εμφανίζονται, δηλαδή 51-58 και 126-134.





Εικόνα 11. Αριστερά τα πλαίσια του πλαστού βίντεο. Δεξιά η αντίστοιχη οπτική ροή.

Η πλαστογραφία που έχει εφαρμοστεί παρατηρείται στα πλαίσια 67-80 και 110-123 μέσω αντιγραφής και αναδιάταξης τμήματος μεταξύ των δύο ομάδων πλαισίων. Στο σημείο αυτό πρέπει να επισημανθεί πως ο αλγόριθμος δεν έχει τη δυνατότητα να δείξει σε ποια από τις δύο ομάδες έχει προκύψει η πλαστογραφία αλλά μόνο την ύπαρξή της. Ανεξαρτήτως αυτού, η αναγνώριση παραβίασης της ακεραιότητας του βίντεο είναι εφικτή και επαναλαμβάνεται η παρατήρηση περί της σχέσης των αποστάσεων των κορυφών της οπτικής ροής στα πλαίσια όπου η παραβίαση αναγνωρίζεται.



Διάγραμμα 6. Ομοιότητες της οπτικής ροής για πραγματική πλαστογραφία και περίπτωση λάθους εκτίμησης.

4.2.6 Αυτοκίνητα ίδιου τύπου σε διαγώνια κίνηση

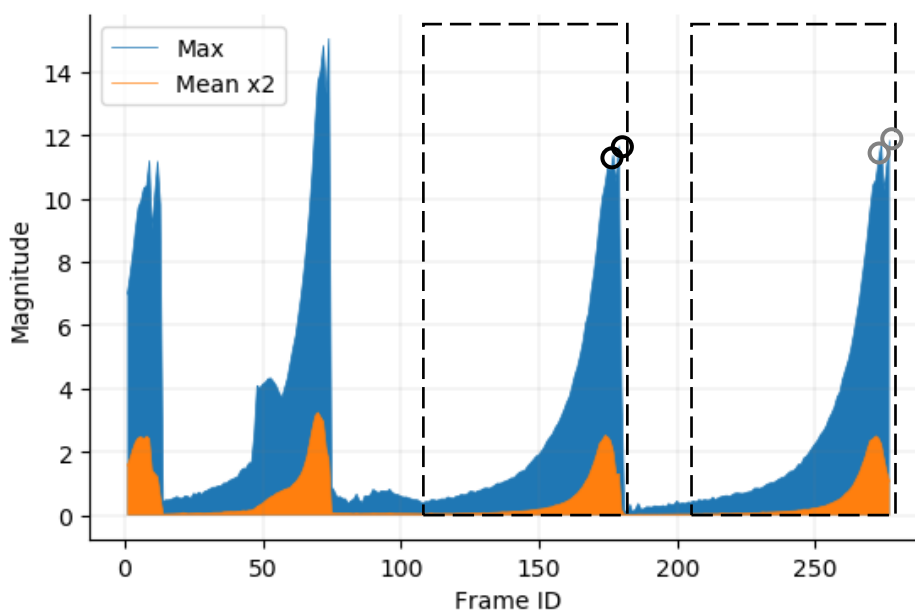
Το τελευταίο βίντεο του συγκεκριμένου συνόλου παρουσιάζει αυτοκίνητα με κατακόρυφη και μερικώς διαγώνια κίνηση σε δρόμο δύο λωρίδων. Η πλαστογραφία εφαρμόστηκε στο αυτοκίνητο του πλαισίου 256 (Εικόνα 12) με αντιγραφή και

αναδιάταξη. Το αυτοκίνητο του πλαισίου 159 ανήκει στην ομάδα των πλαστών πλαισίων και αναγνωρίζεται σωστά από τον αλγόριθμο (Διάγραμμα 7). Ένα πρόβλημα που προκύπτει είναι πως ο αλγόριθμος στη συγκεκριμένη περίπτωση παράγει δεδομένα που μπορούν να οδηγήσουν σε ψευδώς σωστά συμπεράσματα. Τα αυτοκίνητα των πλαισίων 4 και 58 έχουν όμοια κίνηση και χρώμα (ασημί και άσπρο) και το αποτέλεσμα είναι να χαρακτηρίζονται από όμοια οπτική ροή. Τέτοιες ομοιότητες στην ταχύτητα, κατεύθυνση και αντανάκλαση φωτεινών πηγών μπορούν να συμβούν συχνά σε ηλεκτρικούς σηματοδότες ή περιοχές με χαμηλό όριο ταχύτητας και περιορισμένο χώρο κίνησης.



Εικόνα 12. Αριστερά τα πλαίσια του πλαστού βίντεο. Δεξιά η αντίστοιχη οπτική ροή. Παρατηρείται ομοιότητα καθώς τα I.X του αυθεντικού βίντεο είναι όμοια.

Ένα επιπλέον πρόβλημα που προκύπτει είναι ότι η παράλληλη κίνηση του αυτοκινήτου της δεξιάς λωρίδας μεταξύ των πλαισίων 46-116 επηρεάζει τις μέγιστες τιμές της οπτικής ροής. Αν η πλαστογραφία είχε εφαρμοστεί μέσω περικοπής και αντιγραφής πλαισίων σε αυτό το σημείο του βίντεο θα ήταν πολύ δύσκολο να παρατηρηθεί από τον αλγόριθμο. Μία πιθανή λύση στο πρόβλημα είναι η χρήση μάσκας (Bidokhti & Ghaemmaghami, 2015) για το διαχωρισμό των πλαισίων σε περιοχές ενδιαφέροντος όπως λωρίδες κυκλοφορίας.



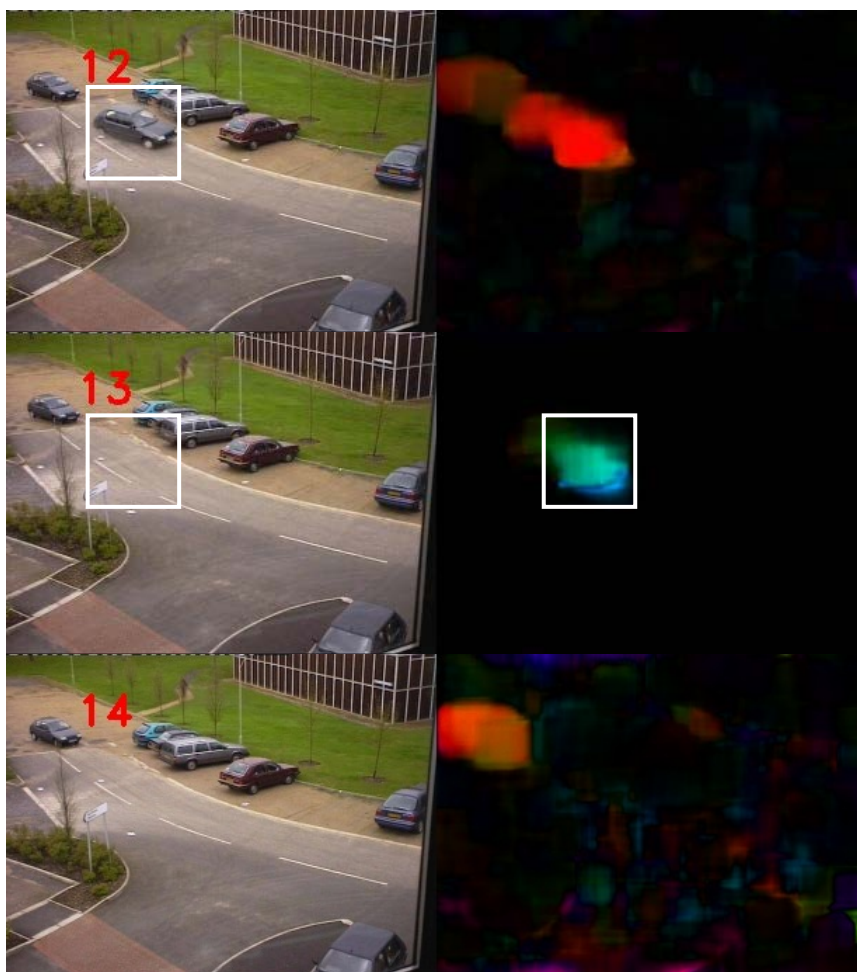
Διάγραμμα 7. Ομοιότητες στην οπτική ροή για το αυθεντικό και το πλαστό αντικείμενο. Αυξημένη πιθανότητα για false-negative καθώς οι κορυφές απέχουν ίση απόσταση για πλαστά και αυθεντικά πλαίσια.

Μία δεύτερη πιθανή λύση είναι η τμηματοποίηση των πλαισίων σε υπερ-εικονοστοιχεία μέσω αλγορίθμων όπως ο SLIC, χρήση μάσκας για την προσπέλαση των υπερ-εικονοστοιχείων και υπολογισμός της οπτικής ροής σε κάθε ένα ξεχωριστά. Κάτι τέτοιο αυξάνει την πολυπλοκότητα του αλγορίθμου και καθιστά δύσκολη την εφαρμογή του σε συσκευές του IoT.

4.2.7 Αφαίρεση πλαισίου σε βίντεο χώρου στάθμευσης

Τα επόμενα πειράματα ανήκουν στο σύνολο του CVIP group και περιλαμβάνουν βίντεο από χώρο στάθμευσης. Οι παραβιάσεις που έχουν υποστεί αφορούν διαγραφή ή αντιγραφή και εισαγωγή περιοχών πλαισίων (spatio-temporal tampering). Το πρώτο βίντεο που ελέγχθηκε έχει υποστεί διαγραφή του μπροστινού αυτοκινήτου όπως έχει

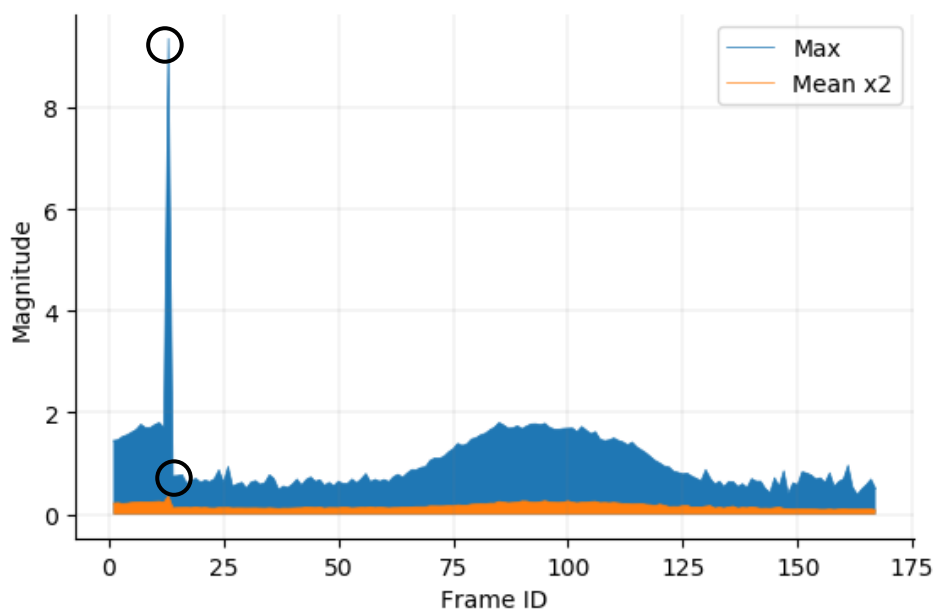
σημειωθεί στο πλαίσιο 12 (Εικόνα 13). Η διαγραφή του αυτοκινήτου εφαρμόζεται στα πλαίσια 12-13 μέσω του εργαλείου που έχει σχεδιαστεί για την παραγωγή συνόλων πλαστογραφημένων βίντεο (Ardizzone & Mazzola, 2015).



Εικόνα 13. Αφαίρεση πλαισίου και απότομη απώλεια οπτικής ροής.

Παρατηρείται πως ενώ υπάρχει απώλεια της οπτικής ροής για το πίσω αυτοκίνητο, η οπτική ροή για το διαγραμμένο έχει αυξηθεί απότομα στο πλαίσιο 13, κάτι που δεν είναι ορατό στο βίντεο αλλά μπορεί να εξηγηθεί μέσω του αλγορίθμου. Ένας λόγος που αυτό προκύπτει είναι γιατί το έδαφος έχει πολύ μεγάλη διαφορά φωτεινότητας από αυτή του αυτοκινήτου που διαγράφεται με αποτέλεσμα να υπάρχει απότομη μετακίνηση των περιοχών φωτεινότητας. Δηλαδή, η μετατόπιση των φωτεινών περιοχών μεταξύ των δύο πλαισίων είναι τέτοια ώστε οι τιμές της ροής για τα πλαίσια 12-13 να παρουσιάζουν μειωμένη συσχέτιση. Η απώλεια της οπτικής ροής για το πίσω αυτοκίνητο μπορεί να οφείλεται σε σφάλμα που εισάγεται μέσω της διπλής συμπίεσης για το συγκεκριμένο πλαίσιο. Εκτελώντας το ίδιο πείραμα με διαφορετικές παραμέτρους για τα επίπεδα της πυραμίδας Gauss και την τιμή του window

παρατηρήθηκαν τα ίδια αποτελέσματα. Μειώνοντας την τιμή του window κάτω από 9 παρατηρήθηκε απώλεια της οπτικής ροής σε όλο το βίντεο με μοναδικό σημείο αύξησής της το πλαίσιο 13 όπου προκύπτει η διαγραφή. Αλλαγές στον αριθμό επαναλήψεων του αλγορίθμου ανά επίπεδο της πυραμίδας Gauss παράγουν επίσης τα ίδια αποτελέσματα. Κάτι τέτοιο επιβεβαιώνει την αρχική υπόθεση ώστε να μην ανάγεται σε λάθος του αλγορίθμου. Οι δύο περιπτώσεις παρουσιάζονται στατιστικά στο Διάγραμμα 8.

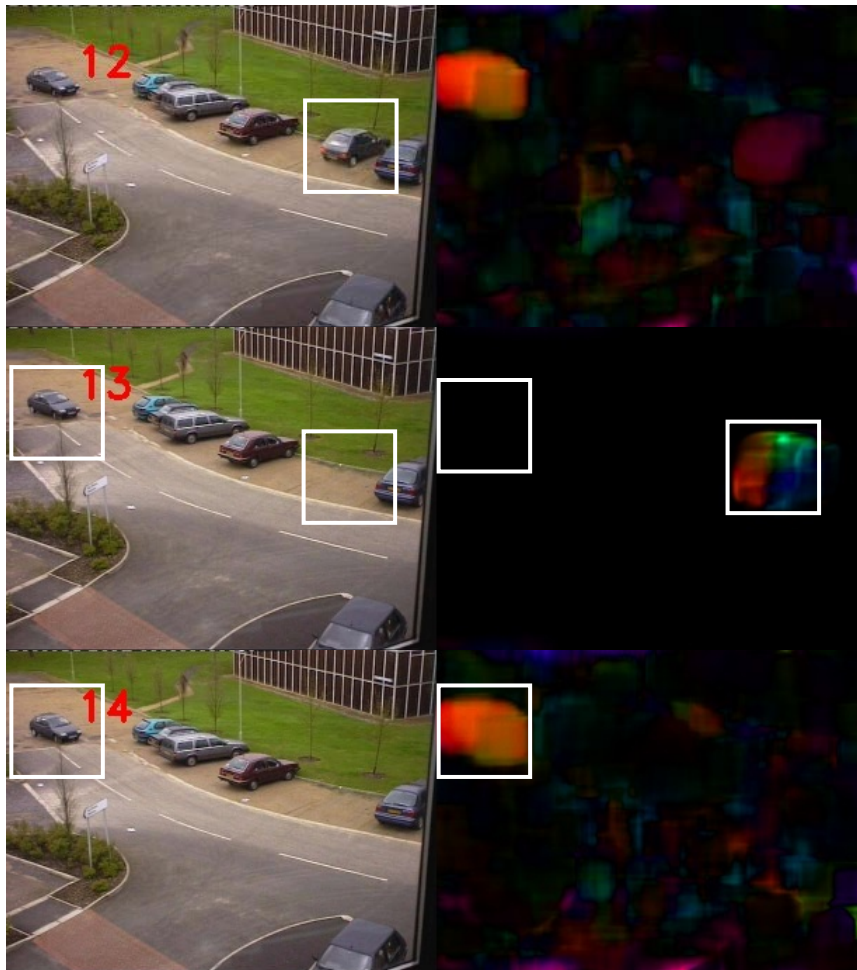


Διάγραμμα 8. Στατιστική απόδοση της άμεσης απώλειας οπτικής ροής με τη διαγραφή πλαισίου.

4.2.8 Χώρος στάθμευσης με χαμηλή οπτική ροή

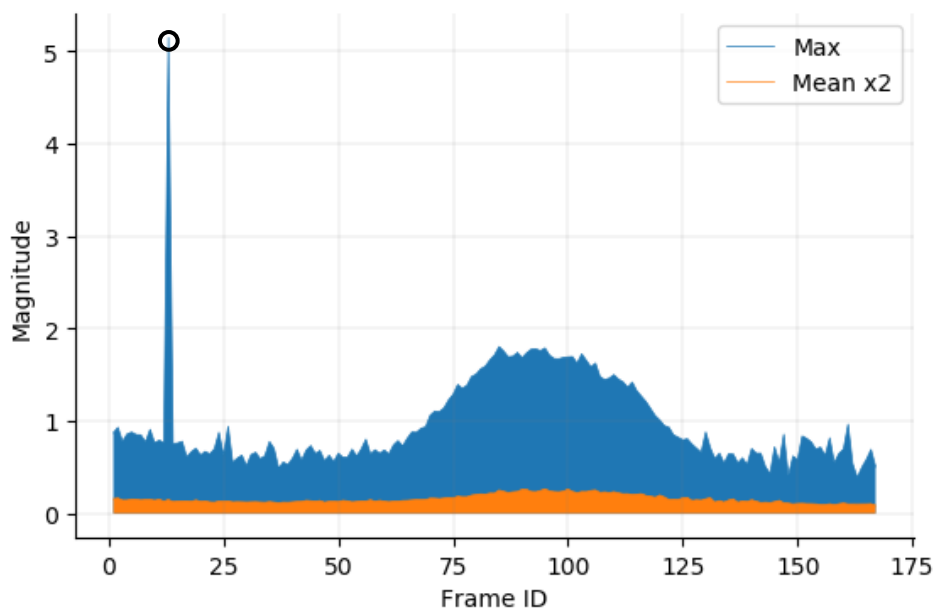
Ένα εξίσου ενδιαφέρον παράδειγμα αποτέλεσε το επόμενο βίντεο καθώς επιβεβαιώνει τα αποτελέσματα του προηγούμενου πειράματος και μπορεί να αναδείξει υπό συνθήκες το αν η πλαστογραφία εφαρμόστηκε σε χωρικό και όχι χρονικό επίπεδο. Αρχικά οι τιμές της οπτικής ροής για το αυτοκίνητο που σταθμεύει είναι αναμενόμενα χαμηλές λόγω της χαμηλής ταχύτητας κατά της στάθμευση. Η διαγραφή του αυτοκινήτου γίνεται με τη χρήση τμήματος από πλαίσιο όπου η θέση στάθμευσης είναι κενή. Επομένως οι περιοχές αυξημένης φωτεινότητας για το συγκεκριμένο σημείο παρουσιάζουν μικρή συσχέτιση μεταξύ των πλαισίων 12 και 13. Όπως φαίνεται στην Εικόνα 14, το αυτοκίνητο δε φωτίζεται πλήρως λόγω της θέσης της πηγής φωτός που δημιουργεί σκίαση. Αντιθέτως δεν υπάρχει κάτι να δημιουργήσει σκίαση στην κενή θέση στάθμευσης του πλαισίου 13. Επομένως η απότομη μετατόπιση των περιοχών

φωτεινότητας οδηγεί στην απότομη αύξηση της οπτικής ροής στο συγκεκριμένο σημείο.



Εικόνα 14. Απότομες μεταβολές φωτεινότητας σπάνε την οπτική ροή.

Η οπτική ροή «σπάει» για το αυτοκίνητο στα αριστερά του βίντεο με τρόπο όμοιο με το προηγούμενο πείραμα. Εκτελώντας τον αλγόριθμο πάνω στο συγκεκριμένο βίντεο με διαφορετικές παραμέτρους λαμβάνουμε τα ίδια αποτελέσματα. Επομένως η απώλεια οπτικής ροής για το δεξί αυτοκίνητο μπορεί επίσης να αποδοθεί σε σφάλμα διπλής συμπίεσης. Η στατιστική προβολή των τιμών της οπτικής ροής φαίνεται στο Διάγραμμα 9.



Διάγραμμα 9. Στατιστική απόδοσης της απότομης απώλειας οπτικής ροής.

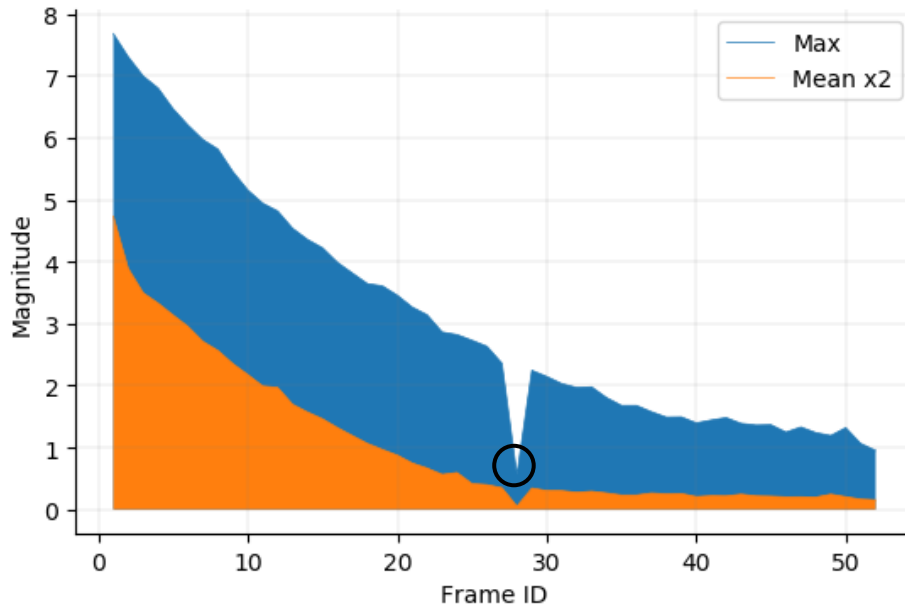
4.2.9 Χώρος στάθμευσης με αυξομείωση των τιμών φωτεινότητας

Το τελευταίο βίντεο που ελέγχθηκε παρουσιάζει δύο Βαν σε διαγώνια κίνηση. Το σημειωμένο Βαν στη Εικόνα 15 εξαφανίζεται με τρόπο που φαίνεται ομαλός στο ανθρώπινο μάτι αλλά παρουσιάζεται απότομος βάσει της οπτικής ροής. Κάτι τέτοιο παρατηρείται στα πλαίσια 24-25 όπου τα δύο αυτοκίνητα χαρακτηρίζονται από διαφορετική οπτική ροή ενώ φαινομενικά έχουν την ίδια ταχύτητα. Η πλαστογραφία είναι ορατή στα πλαίσια 28-29 όπου παρατηρούνται δύο εκτροπές στην οπτική ροή. Αρχικά, η γενική φωτεινότητα του βίντεο αυξάνεται και παρατηρείται εισαγωγή κίνησης στο κάτω δεξιά σημείο του πλαισίου 28 κάτι που μπορεί να ερμηνευτεί ως υπόλειμμα αντιστάθμισης κίνησης MCEA. Επιπρόσθετα, στο ίδιο πλαίσιο υπάρχει απώλεια της οπτικής ροής για το πάνω αριστερά Βαν η οποία είναι όμοια με τις περιπτώσεις των δύο προηγούμενων πειραμάτων και επανεμφανίζεται στο πλαίσιο 29.

Η κατανομή των τιμών μεγέθους της οπτικής ροής για το συγκεκριμένο βίντεο φαίνεται στο Διάγραμμα 10. Η αυξομείωση στο πλαίσιο 28 είναι απότομη και δε συσχετίζεται με την κατανομή ανά τα υπόλοιπα πλαίσια. Επομένως, όπως και στα προηγούμενα πειράματα, η συγκεκριμένη μέθοδος πλαστογραφίας είναι εύκολη στην αναγνώριση μέσω του αλγορίθμου.



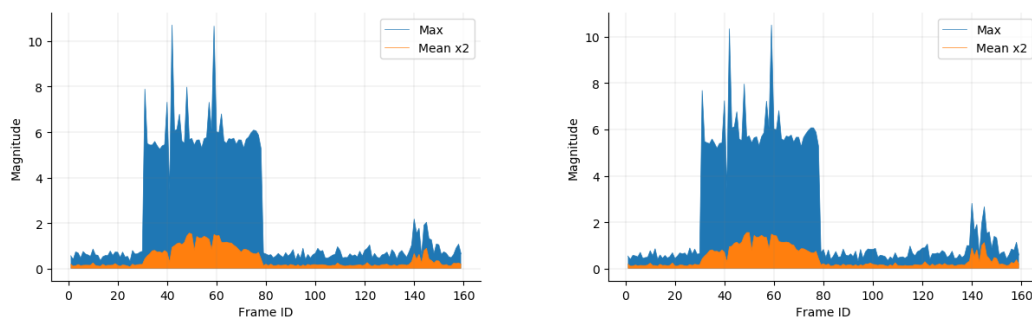
Εικόνα 15. Η οπτική ροή περιορίζεται απότομα για το αριστερό Βαν μεταξύ των πλαισίων 24 και 25. Οι αυξομειώσεις της φωτεινότητας με την αφαίρεση του Βαν προδίδουν την αλλοίωση στα πλαίσια 28 και 29.

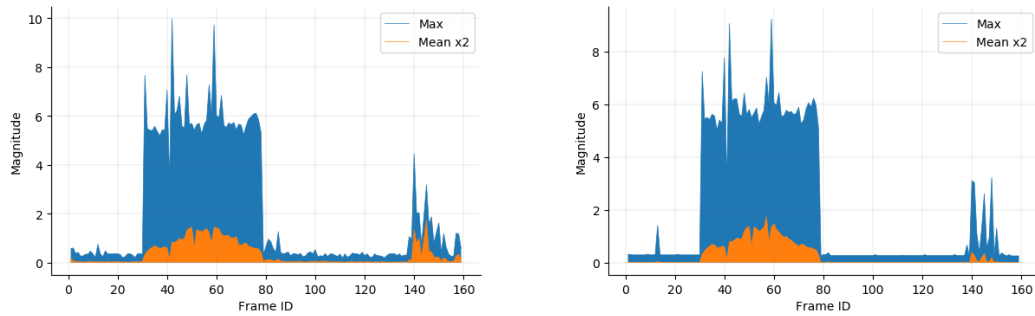


Διάγραμμα 10. Στατιστική απεικόνιση της πλαστογραφίας που προκύπτει στα πλαίσια 28-29.

4.3 Έλεγχος αντοχής σε συμπίεση

Ένα βασικό χαρακτηριστικό των αλγορίθμων αναγνώρισης και ανίχνευσης πλαστογραφίας είναι η ανθεκτικότητά τους σε συμπίεση MPEG-x. Για το εργαλείο της παρούσας μεταπτυχιακής διατριβής χρησιμοποιήθηκαν συμπιεσμένες εκδόσεις σε H.264 των βίντεο από το σύνολο του REWIND project. Ένα αντιπροσωπευτικό πείραμα παρουσιάζεται στο Σχήμα 4. Παρατηρείται πως η σταδιακή εφαρμογή απωλεστικής (lossy) συμπίεσης έχει περιορισμένη επιρροή στα αποτελέσματα του αλγορίθμου. Οι μέγιστες τιμές της οπτικής τιμής μειώνονται ενώ οι κορυφές της σταδιακά εξομαλύνονται αλλά χαρακτηριστικά όπως οι αποστάσεις τους παραμένουν απaráλλακτα και ανιχνεύσιμα. Υποθέτουμε πως με τη χρήση έντονης συμπίεσης ο αλγόριθμος θα παράγει μη αξιόπιστα αποτελέσματα αλλά παρουσιάζει στιβαρότητα έως ένα βαθμό.





Σχήμα 4. Αλλαγές στην οπτική ροή με αυξανόμενη συμπίεση του βίντεο ξεκινώντας από το πάνω αριστερά διάγραμμα.

4.4 Ανάλυση αποτελεσμάτων και βελτιστοποιήσεις

Τα πειράματα που παρουσιάστηκαν είναι στο σύνολό τους 9. Στην πράξη, ο αριθμός πειραμάτων είναι πολύ μεγαλύτερος καθώς για κάθε βίντεο που αναλύθηκε εφαρμόστηκαν διαφορετικές τιμές στις παραμέτρους του αλγορίθμου Gunnar Farneback. Για τα τελικά αποτελέσματα της μεταπτυχιακής διατριβής χρησιμοποιήθηκαν οι τιμές που θεωρήθηκαν βέλτιστες ώστε να παρέχουν τη μέγιστη δυνατή ακεραιότητα και ακρίβεια συνδυάζοντας την καλύτερη δυνατή ταχύτητα ανάλυσης. Ο Πίνακας 2 παρουσιάζει συνοπτικά τα αποτελέσματα της εκτέλεσης του εργαλείου της μεταπτυχιακής διατριβής. Για την ακρίβεια της αναγνώρισης αντιγραφής αντικειμένων ή ολόκληρων πλαισίων χρησιμοποιήθηκαν οι ποιοτικές τιμές «Καλή», «Μέτρια» και «Περιορισμένη». Οι ίδιες τιμές χρησιμοποιήθηκαν για την περιγραφή της πιθανότητας να οδηγηθούμε σε ψευδώς σωστά αποτελέσματα.

Βίντεο	Τύπος Πλαστογραφίας	Ακρίβεια	Ακεραιότητα
4.2.1	Copy-Move	Καλή	Καλή
4.2.2	Copy-Move	Καλή	Καλή
4.2.3	Copy-Move	Καλή	Καλή
4.2.4	Copy-Move	Καλή	Καλή
4.2.5	Copy-Move	Καλή	Μέτρια

4.2.6	Copy-Move	Καλή	Περιορισμένη
4.2.7	Insert/Delete	Καλή	Καλή
4.2.8	Delete	Καλή	Καλή
4.2.9	Delete	Καλή	Καλή

Πίνακας 2. Προσδιορισμός της ακρίβειας και ακεραιότητας του αλγορίθμου

Ο αλγόριθμος παρουσιάζει στιβαρότητα ως προς την αναγνώριση πλαστογραφίας σε χωρικό και χρονικό επίπεδο και οι τιμές του μεγέθους της οπτικής ροής αποτελούν μία ικανή βάση για στατιστική ανάλυση. Μία μέθοδος για αυτοματισμό της ανίχνευσης αντιγραφής με αναδιάταξη είναι ο έλεγχος αποστάσεων μεταξύ των κορυφών. Έχει παρατηρηθεί πως οι αποστάσεις των κορυφών της οπτικής ροής για τα αντικείμενα που έχουν υποστεί αντιγραφή και αναδιάταξη είναι ίσες με διαφορά ενός πλαισίου. Δηλαδή αν οι κορυφές της οπτικής ροής είναι A, B, C για ένα αντικείμενο και A', B', C' αντίστοιχα για τα πλαίσια όπου εφαρμόζεται η αντιγραφή του, τότε η σχέση των αποστάσεων μπορεί να περιγραφεί ως εξής:

$$Distance_{AB} = \begin{cases} Distance_{A'B'} + 1 \\ \text{ή} \\ Distance_{A'B'} \\ \text{ή} \\ Distance_{A'B'} - 1 \end{cases}$$

Με βάση την παραπάνω παρατήρηση μπορεί να θεωρηθεί ότι αν δοσμένου ενός ορίου (threshold), αν η διαφορά των αποστάσεων μεταξύ 2 κορυφών άνω του ορίου ανήκει στο διάστημα [-1,0,1], τότε το βίντεο είναι ύποπτο για πλαστογραφία με αντιγραφή και αναδιάταξη πλαισίων και πρέπει να ελεγχθεί.

Βίντεο	Ανάλυση	Χρόνος του βίντεο	Πλαίσια ανά δευτερόλεπτο	Χρόνος εκτέλεσης
4.2.1	320x240	18s	25	18,34s
4.2.2	320x240	12s	25	16,7s

4.2.3	320x240	12s	25	16,7s
4.2.4	320x240	16s	25	22,41s
4.2.5	320x240	10s	25	15,9s
4.2.6	320x240	22s	25	29.6s
4.2.7	960x540	10s	30	16,09s
4.2.8	960x540	10,36s	30	16,05s
4.2.9	960x540	3,92s	25	8,59s

Πίνακας 3. Χρόνοι εκτέλεσης ανά βίντεο με βέλτιστες και ταχύτερες ρυθμίσεις

Στα βίντεο που ελέγχθηκαν υπάρχουν δύο περιπτώσεις όπου εισάγεται η πιθανότητα για ψευδώς σωστά αποτελέσματα. Όπως έχει προαναφερθεί, μία βασική μεταβλητή που προκαλεί διαφοροποιήσεις ή ομοιότητες στην οπτική ροή είναι η επιφάνεια των κινούμενων αντικειμένων και των στατικών περιοχών. Στο βίντεο της υποενότητας 4.2.6 τα αυτοκίνητα παρουσιάζουν ομοιότητες στην οπτική ροή γιατί έχουν όμοιες, λείες επιφάνειες και αντανακλούν τη φωτεινή πηγή με ανάλογο τρόπο. Επίσης, τα αυτοκίνητα κινούνται με την ίδια κατεύθυνση και ταχύτητα. Επομένως η αναγνώριση πλαστογραφίας σε σχετικές περιπτώσεις βίντεο είναι δύσκολη έως αδύνατη με την αποκλειστική χρήση των χαρακτηριστικών της οπτικής ροής. Η χρήση μηχανικής μάθησης μπορεί να ενισχύσει την ακεραιότητα των αποτελεσμάτων μέσω ανάλυσης των χαρακτηριστικών κάθε επιφάνειας ώστε να υπάρχει διαχωρισμός των αντικειμένων και χρήση της παραπάνω εξίσωσης μόνο όταν παρουσιάζουν ομοιογένεια. Μία δεύτερη παρατήρηση είναι πως σε βίντεο όπου υπάρχει πλαστογραφία σε χωρικό επίπεδο, η οπτική ροή μπορεί να επηρεαστεί από άλλα αντικείμενα που κινούνται εντός των παραβιασμένων πλαισίων. Πολλαπλές περιπτώσεις περικοπής και αναδιάταξης μπορούν να μειώσουν περαιτέρω την αξιοπιστία του αλγορίθμου.

Για τη σύγκριση των αποτελεσμάτων με τις σχετικές έρευνες που έχουν μελετηθεί χρησιμοποιήθηκε η βιβλιογραφική ανασκόπηση των (Shelke & Kasana, 2020). Η έρευνά τους παραθέτει μία συνοπτική καταγραφή των μειονεκτημάτων των σχετικών ερευνών βάσει των οποίων οι μέθοδοι των (Chao, Jiang, & Tanfeng, 2012), (Wang, Wang, Jiang, & Sun, 2013) και (Bidokhti & Ghaemmaghani, 2015) παρουσιάζουν μειωμένη αξιοπιστία σε περιπτώσεις όπου υπάρχουν διαφοροποιήσεις στο μήκος των GOP ή τα παραβιασμένα GOP έχουν αφαιρεθεί ή αναδιαταχθεί πλήρως κατά την πλαστογραφία. Συγκριτικά, η μέθοδος που προτάθηκε στην παρούσα μεταπτυχιακή διατριβή δεν περιορίζεται σε χαρακτηριστικά της συμπίεσης MPEG και επομένως δε βασίζεται στην κατανομή των I,P και B πλαισίων. Ως επιπλέον μειονέκτημα των σχετικών ερευνών, το οποίο παρατηρείται ομοίως στο εργαλείο της παρούσας έρευνας, αναφέρεται η αδυναμία στην ανίχνευση πλαστογραφίας σε βίντεο όπου το φόντο είναι κινούμενο. Οι (Shelke & Kasana, 2020) έκριναν πως η μέθοδος των (Chao, Jiang, & Tanfeng, 2012) χρειάζεται βελτίωση ως προς την ανίχνευση αφαίρεσης πλαισίων. Συγκριτικά η παρούσα έρευνα παρουσιάζει αξιοπιστία στην ανίχνευση διαγραφής πλαισίων και περιοχών αυτών αλλά έχει το μειονέκτημα της ανάγκης για χρήση διαφορετικών παραμέτρων για τον αλγόριθμο του Farneback. Συγκεκριμένα, κατά τη διεξαγωγή πειραμάτων απαιτήθηκε συχνά η χρήση διαφορετικού μεγέθους της παραμέτρου window, κάτι που δυσκολεύει την αυτοματοποίηση του εργαλείου. Η χρήση μάσκας για το διαχωρισμό των πλαισίων σε περιοχές ενδιαφέροντος (Bidokhti & Ghaemmaghani, 2015) θεωρείται κρίσιμη για την επίτευξη της αξιοπιστίας της μεθόδου τους και αποτελεί μία αποτελεσματική μέθοδο ανίχνευση στην περίπτωση όπου η πλαστογραφία έχει πραγματοποιηθεί σε χωρικό και όχι μόνο χρονικό επίπεδο. Αν και το εργαλείο της παρούσας έρευνας δεν εξαρτάται από τα GOP όπως η μέθοδος των (Bidokhti & Ghaemmaghani, 2015), η χρήση μάσκας είναι μία μέθοδος που μπορεί να το ενισχύσει ως προς τον περιορισμό του φαινομένου επηρεασμού της οπτικής ροής από πολλαπλά κινούμενα αντικείμενα σε ένα βίντεο. Για παράδειγμα, σε ένα βίντεο δρόμου δύο λωρίδων η μάσκα είναι επαρκής για το διαχωρισμό των πλαισίων σε περιοχές ελέγχου ώστε ο αλγόριθμος που παρουσιάστηκε να εφαρμοστεί ξεχωριστά σε κάθε μία αλλά η ανάλυση των δεδομένων της οπτικής ροής να γίνει συλλογικά για τις δύο περιοχές. Παράλληλα, οι δείκτες των (Wang, Wang, Jiang, & Sun, 2013) αποτελούν μία καλή λύση για την ομαλοποίηση των τιμών του μεγέθους της οπτικής ροής για την ανίχνευση διαγραφής ή προσθήκης πλαισίων. Συνοπτικά η μέθοδος που έχει προταθεί στην παρούσα μεταπτυχιακή διατριβή παρουσιάζει πλεονεκτήματα σε σχέση με τις

αντίστοιχες έρευνες που αναλύθηκαν αλλά μπορεί να εκμεταλλευτεί τεχνικές τις οποίες έχουν προτείνει για την αυτοματοποίηση της ανίχνευσης πλαστογραφίας και την αύξηση της αξιοπιστίας του εργαλείου.

Κεφάλαιο 5

Συμπεράσματα & Επίλογος

Στόχος της παρούσας μεταπτυχιακής διατριβής υπήρξε το να προτείνει μία νέα μέθοδο αναγνώρισης πλαστογραφίας σε βίντεο και να παρουσιάσει την υλοποίηση ενός εργαλείου ικανού για την εφαρμογή της. Για την επίτευξη του σχεδιασμού και της ανάπτυξης του κρίθηκε απαραίτητη η μελέτη και ανάλυση του γενικού ερευνητικού πεδίου της δικανικής για την ανίχνευση πλαστογραφίας σε βίντεο. Προηγήθηκε η βιβλιογραφική ανασκόπηση ερευνών που ασκούν κριτική σε πρόσφατες και παλαιότερες μεθόδους ανίχνευσης με στόχο τον καθορισμό της ερευνητικής βάσης της μεταπτυχιακής διατριβής. Η μέθοδος που σχεδιάστηκε ανήκει στην ομάδα τεχνικών ανίχνευσης με βάση τα χαρακτηριστικά κίνησης και στηρίχθηκε στην ανάλυση της οπτικής ροής. Σε αντίθεση με τις μέχρι τώρα έρευνες που χρησιμοποιούν τη μέθοδο των Lucas-Kanade, η παρούσα μεταπτυχιακή διατριβή πρότεινε και εφάρμοσε τον αλγόριθμο υπολογισμού πυκνής οπτικής ροής του Gunnar Farneback. Οι δύο λόγοι για την επιλογή του υπήρξαν ότι κατά τη γνώση του συγγραφέα η συγκεκριμένη μέθοδος δεν έχει χρησιμοποιηθεί στην αναγνώριση βίντεο και ότι σχετική έρευνα έχει αποδείξει την αξιοπιστία του αλγόριθμου στην ακρίβεια του υπολογισμού της οπτικής ροής. Μία δεύτερη πρόταση της παρούσας έρευνας υπήρξε η εφαρμογή του εργαλείου σε συσκευή που συχνά συναντάται στο αυξανόμενο διαδίκτυο των πραγμάτων (IoT). Το εργαλείο υλοποιήθηκε με τη χρήση της βιβλιοθήκης OpenCV σε μία συσκευή χαμηλού κόστους

και περιορισμένων υπολογιστικών επιδόσεων Raspberry Pi 4B. Η εκτέλεσή του μέσα από ένα σύνολο πειραμάτων απέδειξε αρχικά ότι η συγκεκριμένη συσκευή είναι αρκετά ισχυρή ώστε να επιτύχει την ομαλή εκτέλεση του εργαλείου χωρίς να παρουσιάσει αστάθειες στη λειτουργία της. Δεύτερον, ο αλγόριθμος είναι ικανός να αναγνωρίσει την κακόβουλη εισαγωγή, διαγραφή και αντιγραφή ομάδων πλαισίων και περιοχών εντός αυτών. Η ακρίβειά του κρίθηκε καλή και η ακεραιότητά του ανθεκτική στο θόρυβο και στην έως ένα βαθμό συμπίεση των βίντεο. Τέλος, προτάθηκαν μεθοδολογίες αυτόματης ανίχνευσης της πλαστογραφίας με τη χρήση χαρακτηριστικών που εξήχθησαν από την εκτέλεση του αλγορίθμου και παρουσίασαν έντονη συσχέτιση μεταξύ των πειραμάτων εκτέλεσης.

Ως μελλοντική έρευνα προτείνεται η ανάπτυξη των μεθόδων αυτόματης ανίχνευσης με τη χρήση μάσκας για διαχωρισμό των πλαισίων σε περιοχές ενδιαφέροντος όμοια με την έρευνα των (Bidokhti & Ghaemmaghami, 2015) και η εφαρμογή των δεικτών που έχουν χρησιμοποιήσει οι (Wang, Wang, Jiang, & Sun, 2013). Παράλληλα, συνιστάται η εφαρμογή τεχνικών αναγνώρισης χαρακτηριστικών ομοιογένειας και μοντέλων μηχανικής μάθησης καθώς μπορεί να περιορίσει την πιθανότητα εμφάνισης ψευδώς σωστών αποτελεσμάτων. Με το σχεδιασμό των παραπάνω προτείνεται ως τελικό στάδιο η πειραματική εφαρμογή του εργαλείου σε άλλες συσκευές του διαδικτύου των πραγμάτων και η διεξαγωγή πειραμάτων σε νέα σύνολα βίντεο ώστε να επιβεβαιωθεί εκ νέου η σημαντικότητα των παραπάνω προτάσεων έρευνας.

Βιβλιογραφία

Antoniadis, A., Sklavos, N., & Kavum, E. B. (2020). An Efficient Implementation of A Delay-Based PUF Construction. *Design, Automation and Test in Europe Conference (DATE'20)*. Paris, France.

Ardizzone, E., & Mazzola, G. (2015). A Tool to Support the Creation of Datasets of Tampered Videos. *International Conference on Image Analysis and Processing, 9280*, pp. 665-675.

Bestagini, P., Milani, S., Tagliasacchi, M., & Tubaro, S. (2013). Local tampering detection in video sequences. *2013 IEEE International Workshop on Multimedia Signal Processing (MMSP)*, (pp. 488-493). Pula, Italy.

Bidokhti, A., & Ghaemmaghami, S. (2015). Detection of regional copy/move forgery in MPEG videos using optical flow. *The International Symposium on Artificial Intelligence and Signal Processing (AISP)*, (pp. 13-17). Mashhad.

Chao, J., Jiang, X., & Tanfeng, S. (2012). A Novel Video Inter-frame Forgery Model Detection Scheme Based on Optical Flow Consistency. In *The International Workshop on Digital Forensics and Watermarking* (pp. 267-281). Berlin, Heidelberg: Springer.

De Boer, J., & Kalksma, M. (2015). Choosing Between Optical Flow Algorithms for UAV Position Change Measurement.

- Farneback, G. (2003). Two-Frame Motion Estimation Based on Polynomial Expansion (SCIA). In *Image Analysis* (Vol. LNCS 2749, pp. 363-370).
- Fayyaz, M. A., Anjum, A., Ziauddin, S., & Khan, A. (2020). An improved surveillance video forgery detection technique using sensor pattern noise and correlation of noise residues. *Multimedia Tools and Applications*, 79.
- Isa, E., & Sklavos, N. (2017). Smart Home Automation: GSM Security System Design & Implementation. *3rd Pan-Hellinic Conference on Electronics and Telecommunications (PACET'15), Selected Paper, Journal of Engineering Science and Technology Review, vol 10*. Ioannina, Greece.
- Jeong, Y., Moon, K., & Kim, J. (2008). Implementation of real time video watermark embedder based on Haar wavelet transform using FPGA. *Second International Conference on Future Generation Communication and Networking Symposia*, (pp. 63-66). Sanya.
- Johnston, P., & Elyan, E. (2019). A Review of Digital Video Tampering: From Simple Editing to Full Synthesis. *Digital Investigation* 29, pp. 67-81.
- Kalis, R., & Belloum, A. S. (2018). Validating data integrity with blockchain. *2018 IEEE International Conference on Cloud Computing Technology and Science*, (pp. 272-277).
- Kaur, S., & Kushwaha, M. (2018). A Comparative study of various Video Tampering detection methods. *ICSCCC*, (pp. 418-423). Jalandhar.
- Limnaios, S., Sklavos, N., & Koufopavlou, O. (2019). Lightweight Efficient Simeck32/64 Crypto-Core Designs and Implementations, for IoT Security. *27th IFIP/IEEE International Conference On Very Large Scale Integration (VLSI-SoC'19)*. Cusco, Peru.
- Lucas, B. D., & Kanade, T. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision (IJCAI). *Proceedings of Image Understanding Workshop*, (pp. 121-130).
- Panagiotou, P., Sklavos, N., & Zaharakis, I. D. (2018). Design and Implementation of a Privacy Framework, for the Internet of Things (IoT). *21th EUROMICRO Conference on Digital System Design, Architectures, Methods, Tools (DSD'18)*. Prague, Czech Republic.

- Pantopoulou, M., & Sklavos, N. (2020). An FPGA-Implemented Parallel System of Face Recognition, for Digital Forensics Applications. *10th IEEE International Conference on Consumer Technology*. Berlin, Germany.
- Pun, C.-M., & Chung, J.-L. (2018). A Two-stage Localization for Copy-Move Forgery Detection. *Information Sciences*, pp. 33-55.
- Qadir, G., Yahaya, S., & T.S Ho, A. (2012). Surrey University Library for Forensic Analysis (SULFA) of video content. *IET Conference on Image Processing(IPR 2012)*, (pp. 1-6). London.
- Radhakrishnan, R., Sharma, D., & Murthy, V. (2017). A Review On Particle Image Velocimetry And Optical Flow Methods In Riverine Environment. *Vishwacon Proceedings of the 1st International Conference on Recent Trends in Engineering & Technology*. Pune, India.
- Shelke, N., & Kasana, S. (2020). A comprehensive survey on passive techniques for digital video forgery detection. *Multimedia Tools and Applications*, pp. 1-64.
- Shojanazeri, H., Wan Adnan, W. A., & Mumtadzah Syed Ahmad, S. (2013). Video Watermarking Techniques for Copyright protection and Content Authentication. *ISSN 2150-7988*, pp. 652-660.
- Sklavos, N., & Francese, S. (2020). MULTI-modal Imaging of FOREnsic SciEnce Evidence: MULTI-FORESEE Project. *23th EUROMICRO Conference on Digital System Design, Architectures, Methods, Tools (DSD'20)*. Portoroz, Slovenia.
- Sklavos, N., & Souras, P. (2006, January). Economic Models & Approaches in Information Security for Computer Networks. *International Journal of Network Security (IJNS)*, vol. 2, pp. 14-20.
- Sklavos, N., & Toulou, K. (2007). A System-Level Analysis of Power Consumption & Optimizations in 3G Mobile Devices. pp. 225-235.
- Sklavos, N., Chaves, R., Di Natale, G., & Regazzoni, F. (2017). *Hardware Security and Trust*. Springer.
- Sklavos, N., Hübner, M., Göhringer, D., & Kitsos, P. (2013). *System-Level Design Methodologies for Telecommunication*. Springer.

Spanaki, P., & Sklavos, N. (2018). Cloud Computing: Security Issues and Establishing Virtual Cloud Environment via Vagrant to Secure Cloud Hosts. In K. Daimi, *Computer and Network Security Essentials* (pp. 539-553). Springer.

Theodorou, S., & Sklavos, N. (2019). Blockchain-Based Security and Privacy in Smart Cities. In D. B. Rawat, & K. Z. Ghafoor, *Smart Cities Cybersecurity and Privacy* (pp. 21-37). Elsevier Press.

Wang, W., Wang, S., Jiang, X., & Sun, T. (2013). Identifying Video Forgery Process Using Optical Flow. In *Digital-Forensics and Watermarking* (Vol. 8389, pp. 244-257). Berlin, Heidelberg: Springer.

Yatskiv, V., Yatskiv, N., & Bandrivskyi, O. (2019). Proof of Video Integrity Based on Blockchain. *2019 9th International Conference on Advanced Computer Information Technologies*, (pp. 431-434).

Zhao, D.-N., Wang, R.-K., & Lu, Z.-M. (2018). Inter-frame passive-blind forgery detection for video shot based on similarity analysis. *Multimedia Tools and Applications* .

Παράρτημα Α

Κώδικας Python

```
import cv2
import numpy as np
import matplotlib.pyplot as plt
import os
import sys

os.system('clear')
os.system('echo "CMF_Optical_Flow"')
print("Ioannis Memos-Bagkratsas\nOpen University of Cyprus\nDecember 2020\n")

#Initial Checks
enable_run = 0
exit_code = 0
if len(sys.argv) != 17:
    exit_code = 1
    if len(sys.argv) == 2 and sys.argv[1] == '-h':
        exit_code = 2
else:
    if sys.argv[1] == '-run':
        enable_run += 1 if sys.argv[3] == '-p' and float(sys.argv[4]) > 0 and float(sys.argv[4]) < 1 else None
        enable_run += 1 if sys.argv[5] == '-l' and int(sys.argv[6]) >= 1 else None
        enable_run += 1 if sys.argv[7] == '-w' and int(sys.argv[8]) > 0 else None
        enable_run += 1 if sys.argv[9] == '-i' else None
        enable_run += 1 if sys.argv[11] == '-n' and (int(sys.argv[12]) == 5 or int(sys.argv[12]) == 7) else None
```

```

    enable_run +=1 if sys.argv[13]=='-s' and (float(sys.argv[14])==1.1 or float(sys.argv[14])==1.5) else
None
    enable_run +=1 if sys.argv[15]=='-scale' and int(sys.argv[16])>=1 and int(sys.argv[16])<=100 else
None
    else:
        exit_code = 1

#Main Code___
if exit_code == 0 and enable_run == 7:
#___Read the video source
    cap = cv2.VideoCapture(sys.argv[2])
#___Assign cmd arguments to variables
    pyr_s = float(sys.argv[4])
    pyr_l = int(sys.argv[6])
    wnd = int(sys.argv[8])
    pyr_i = int(sys.argv[10])
    poly_n = int(sys.argv[12])
    poly_s = float(sys.argv[14])
    scale_percent = int(sys.argv[16])
#___Initialize fps counter and magnitude lists
    count = 0
    magnitude_max = []
    magnitude_min = []
    magnitude_mean = []
    magnitude_count = []
#___Start the capture
    while True:
        ret, frame1 = cap.read()#Read the first frame
        font = cv2.FONT_HERSHEY_DUPLEX
        if ret:
            print("1st frame inserted")
            width = int(frame1.shape[1]*scale_percent/100)
            height = int(frame1.shape[0]*scale_percent/100)
            new_size = (width,height)
            frame1 = cv2.resize(frame1, new_size)
            prvs = cv2.cvtColor(frame1, cv2.COLOR_BGR2GRAY)
            hsv = np.zeros_like(frame1)
            hsv[...,-1] = 255
            ret, frame2 = cap.read()#Read the second frame
            if ret:
                print("2nd frame inserted\n")
                frame2 = cv2.resize(frame2, new_size)
                count += 1
                nxt = cv2.cvtColor(frame2, cv2.COLOR_BGR2GRAY)
#_____Calculate the optical flow
                flow = cv2.calcOpticalFlowFarneback(prvs, nxt, None, pyr_s, pyr_l, wnd, pyr_i, poly_n,
poly_s, 0)

```



```

    mag, ang = cv2.cartToPolar(flow[...,0], flow[...,1])
    print(np.max(mag))
#_____Update magnitude lists
    magnitude_max = np.append(magnitude_max, np.max(mag))
    magnitude_min = np.append(magnitude_min, np.min(mag))
    magnitude_mean = np.append(magnitude_mean, np.mean(mag)*2)
    magnitude_count = np.append(magnitude_count, count)
#_____Prepare image for output
    hsv[...,0] = ang*180/np.pi/2
    hsv[...,2] = cv2.normalize(mag, None, 0, 255, cv2.NORM_MINMAX)
    rgb = cv2.cvtColor(hsv, cv2.COLOR_HSV2BGR)
    cv2.putText(frame2, str(count), (50,50), font, 1, (0,0,255),2,cv2.LINE_4)
    out_frame = np.concatenate((frame2, rgb), axis = 1)
    cv2.imshow('both', out_frame)#output
else:
    print("Frame not inserted or end of frame sequence!\n")
#_____Create the plot
    plt.xlabel('Frame ID')
    plt.ylabel('Magnitude')
    plt.gca().spines['right'].set_visible(False)
    plt.gca().spines['top'].set_visible(False)
    plt.grid(color='grey', linestyle='-', linewidth=0.25, alpha=0.5)
    plt.plot(magnitude_count, magnitude_max, linewidth = 0.5)
    plt.plot(magnitude_count, magnitude_mean, linewidth = 0.5)
    plt.fill_between(magnitude_count, magnitude_max)
    plt.fill_between(magnitude_count, magnitude_mean)
    plt.legend(['Max', 'Mean x2'], loc = 'upper left')
    plt.show()
    break
    exit_key = cv2.waitKey(30) & 0xFF
    if exit_key == ord('q'):
        break
    cap.release()
elif exit_code==2:
    print("This is a python script to calculate and analyze the optical flow of suspicious videos\n")
    print("Example for running the script:\n")
    print("\tpython cmf_detect_ouc.py -run 'path' -p 0.5 -l 5 -w 50 -i 3 -n 7 -s 1.5 -scale 100\n")
    print("Arguments:\n")
    print("-p | Needs to be float. Defines the downscaling rate of each pyramid level\n")
    print("-l | Needs to be int. Defines the number of the pyramid's levels\n")
    print("-w | Needs to be int. Defines the window size. Bigger is more robust but blurry\n")
    print("-i | Needs to be int. Defines the number of the algorithm iterations per pyramid level\n")
    print("-n | Needs to be int. Defines the neighboring pixels. 5 is faster but 7 is more robust\n")
    print("-s | Needs to be float. The standard deviation for the Gaussian. Use 1.1 if n=5 or 1.5 if n=7\n")
elif exit_code==1:
    print("Undefined or wrong arguments. Exiting!\n")

```

`cv2.destroyAllWindows()`

Παράρτημα Β

Οδηγίες Υλοποίησης

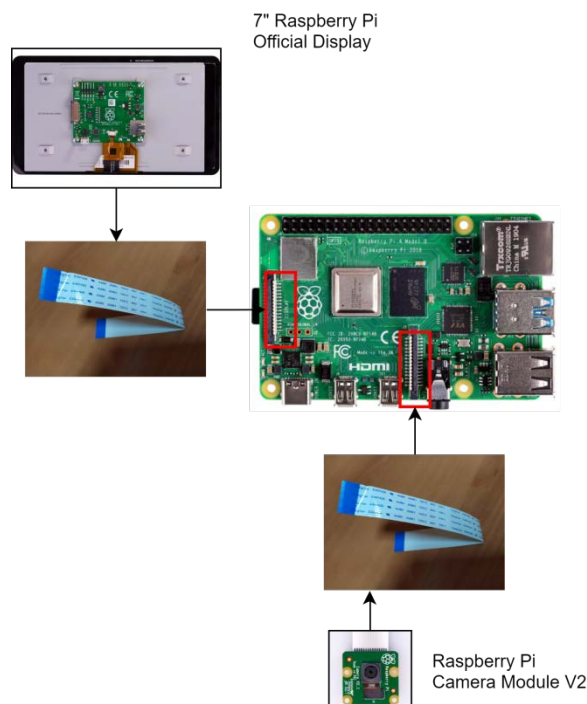
Το συγκεκριμένο παράρτημα αποτελεί έναν αναλυτικό οδηγό της διαδικασίας συναρμολόγησης του υλικού και της εγκατάστασης του λογισμικού όπως έχει εφαρμοστεί στην υλοποίηση της παρούσας μεταπτυχιακής διατριβής. Επιπλέον παρέχονται εναλλακτικές προτάσεις υλικού όπου αυτό καθίσταται εφικτό. Για την καλύτερη κατανόηση παρέχονται στιγμιότυπα οθόνης όπου αυτό κρίθηκε απαραίτητο.

B.1 Απαιτούμενο υλικό και συναρμολόγηση

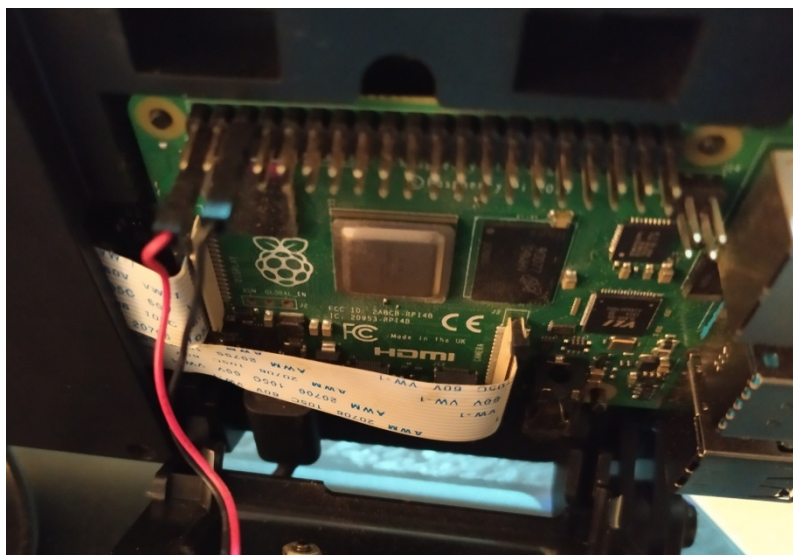
Για την υλοποίηση σε επίπεδο υλικού χρησιμοποιήθηκαν συνοπτικά οι ακόλουθες συσκευές οι οποίες και αναλύονται στη συνέχεια του παραρτήματος.

1. Raspberry Pi 4B με 4GB LPDDR4 RAM.
2. 64GB microSD κάρτα κλάσης ταχύτητας A1.
3. Επίσημη οθόνη αφής Raspberry Pi 7".
4. Επίσημη κάμερα Raspberry Pi Camera v2.
5. Θήκη SmartPi Touch 2 για τη στέγαση του Raspberry Pi 4, της οθόνης 7" και της κάμερας.
6. Δύο καλώδια Ribbon Flex για τη σύνδεση του Raspberry Pi με την οθόνη και την κάμερα.
7. Επίσημος μετασχηματιστής 3A για το μοντέλο Raspberry Pi 4.

Το μοντέλο του Raspberry Pi 4B που χρησιμοποιήθηκε στα πλαίσια της μεταπτυχιακής διατριβής περιλαμβάνει 4GB LPDDR4 RAM αλλά παρέχεται επίσης η δυνατότητα επιλογής 8GB LPDDR4 RAM. Επιπλέον, η επίσημη οθόνη 7" είναι προαιρετική καθώς υποστηρίζεται επίσης η χρήση οθόνης αφής με διασύνδεση mini HDMI ή έως 2 εξωτερικών οθονών μέσω καλωδίων mini HDMI. Παράλληλα, η επιλογή της θήκης για στέγαση του Raspberry Pi 4B και της οθόνης είναι προαιρετική όπως επίσης το μοντέλο που δύναται να χρησιμοποιηθεί. Η θήκη SmartiPi Touch 2 προτιμήθηκε καθώς παρέχει υποδοχή για την επίσημη Pi Camera V2 και επιτρέπει την εναλλακτική τοποθέτηση σε επιφάνεια τοίχου. Ως αποθηκευτικός χώρος προτείνεται μία κάρτα microSD των 64GB με ταχύτητα κλάσης A1 (τουλάχιστον 1500 IOPS ανάγνωση και 500 IOPS εγγραφή) ή A2 (τουλάχιστον 4000 IOPS ανάγνωση και 2000 IOPS εγγραφή). Η επιλογή μεγαλύτερου χώρου αποθήκευσης θεωρείται βέλτιστη για βίντεο μεγάλης διάρκειας ή υψηλής ανάλυσης 4K αλλά αυξάνει το κόστος της υλοποίησης. Τέλος, προτείνεται η χρήση ενός προσαρμογέα microSD σε USB για την εγκατάσταση του λειτουργικού συστήματος σε περίπτωση που η κάρτα microSD δεν έχει ήδη εγκατεστημένο το Raspberry Pi OS ή το NOOBS installer. Για τη σύνδεση της οθόνης 7" και της κάμερας χρησιμοποιούνται δύο καλώδια Ribbon Flex όπως φαίνονται στο Σχήμα 5 και στην Εικόνα 16. Τέλος, ο ανεμιστήρας που παρέχεται με τη θήκη SmartiPi Touch 2 τροφοδοτείται από δύο GPIO pins (Εικόνα 16).



Σχήμα 5. Σύνδεση οθόνης και κάμερας μέσω Ribbon Flex καλωδίων.



Εικόνα 16. Σύνδεση καλωδίων Ribbon Flex στις υποδοχές οθόνης (αριστερά) και κάμερας (δεξιά).

Η τροφοδοσία της συσκευής γίνεται μέσω θήρας USB C όπως φαίνεται στην Εικόνα 17. Για την παράλληλη τροφοδοσία του Raspberry Pi 4B και της οθόνης 7" απαιτείται καλώδιο με υποδοχή USB C το οποίο διακλαδώνεται σε δύο USB C απολήξεις με την αριστερή να τροφοδοτεί την οθόνη και τη δεξιά το Raspberry Pi.



Εικόνα 17. Σύνδεση τροφοδοσίας με καλώδιο δύο διακλαδώσεων σε οθόνη και Raspberry Pi 4B.

B.2 Απαιτούμενο λογισμικό και εγκατάσταση

Η υλοποίηση σε επίπεδο λογισμικού απαιτεί συνοπτικά τα ακόλουθα.

1. Λειτουργικό σύστημα Raspberry Pi OS Desktop.
<https://www.raspberrypi.org/software/operating-systems/>
2. Πρόγραμμα εγκατάστασης σε microSD. Χρησιμοποιήθηκε το Balena Etcher.

<https://www.balena.io/etcher/>

3. Περιβάλλον OpenCV και της γλώσσας προγραμματισμού Python.

<https://opencv.org>

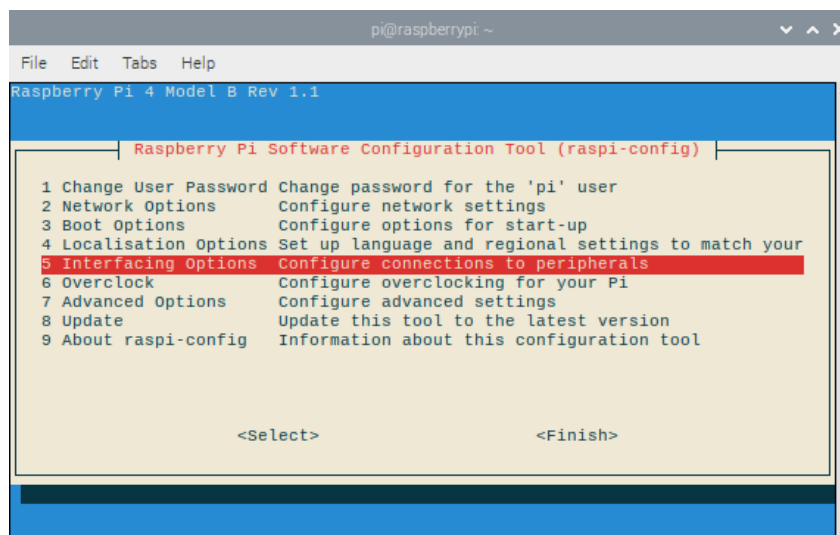
4. Ένα VNC server και VNC client. Χρησιμοποιήθηκε ο RealVNC server που παρέχεται από το Raspberry Pi OS κατά την εγκατάστασή του και ο RealVNC viewer για τη σύνδεση από την απομακρυσμένη συσκευή.

<https://www.realvnc.com/en/connect/download/viewer/>

5. Προαιρετικά το putty.exe για την απομακρυσμένη σύνδεση μέσω πρωτοκόλλου SSH.

<https://www.putty.org>

Το λειτουργικό σύστημα που χρησιμοποιήθηκε είναι το Raspberry Pi OS Desktop. Για την εγκατάστασή του αρχικά πρέπει να πραγματοποιηθεί λήψη από την επίσημη ιστοσελίδα. Στη συνέχεια απαιτείται η εισαγωγή της microSD κάρτας μέσω ενός μετατροπέα σε USB ή SD και η χρήση του προγράμματος Etcher για την εγκατάσταση στη microSD. Όταν ολοκληρωθεί η εγκατάσταση η microSD μπορεί να εισαχθεί στην υποδοχή του Raspberry Pi 4B. Για την εκκίνηση αρκεί η σύνδεση του μετασχηματιστή 3A στην υποδοχή USB C. Η ενεργοποίηση των VNC και SSH servers γίνεται μέσω παραθύρου εντολών. Αρχικά ο χρήστης πληκτρολογεί την εντολή **sudo raspi-config** και μέσω του μενού επιλογών διασύνδεσης (Εικόνα 18) επιλέγει διαδοχικά **VNC>Yes**, **SSH>Yes**. Τέλος, επιλέγει **Enable** και **Finish** για την εφαρμογή των επιλογών του.



Εικόνα 18. Περιβάλλον διαχείρισης υλικού στο Raspberry Pi OS.

Για την απομακρυσμένη σύνδεση και εκτέλεση του εργαλείου σε περιβάλλον γραφικών χρησιμοποιήθηκε το πρόγραμμα RealVNC ενώ για τον προγραμματισμό του εργαλείου

πραγματοποιήθηκε σύνδεση μέσω πρωτοκόλλου SSH με το πρόγραμμα putty.exe και τον VIM editor.

Το περιβάλλον της Python 2 και 3 είναι ήδη εγκατεστημένο στο Raspberry Pi OS. Πριν από την εγκατάσταση της OpenCV προτείνεται αρχικά η λήψη ενημερώσεων μέσω παραθύρου εντολών εκτελώντας διαδοχικά τις εντολές **sudo apt update** και **sudo apt full-upgrade -y**. Το επόμενο απαιτούμενο βήμα είναι η επέκταση του αποθηκευτικού χώρου στη microSD καθώς κατά την εγκατάσταση δεσμεύεται ένα ορισμένο μέγεθος της κάρτας. Η διαδικασία απαιτεί τη εκτέλεση της εντολής **sudo raspi-config** και τις επιλογές **Advanced Options>Expand Filesystem>Enable**. Για την εφαρμογή των αλλαγών απαιτείται η επανεκκίνηση της συσκευής. Η εγκατάσταση της OpenCV γίνεται μέσω παραθύρου εντολών και απαιτεί έναν αριθμό εντολών οι οποίες είναι απαραίτητο να εκτελεστούν διαδοχικά για την ενημέρωση των βιβλιοθηκών του συστήματος ώστε να είναι εφικτή η χρήση της. Οι εντολές καθώς και η επεξήγησή τους είναι οι ακόλουθες και η διαδικτυακή πηγή είναι <https://www.pyimagesearch.com/2018/09/26/install-opencv-4-on-your-raspberry-pi/>.

1. Εγκατάσταση των εξαρτήσεων και του GTK για τη δημιουργία γραφικών περιβαλλόντων διεπαφής (GUI)

1.1.1. sudo apt-get install build-essential cmake unzip pkg-config

1.1.2. sudo apt-get install libjpeg-dev libpng-dev libtiff-dev

1.1.3. sudo apt-get install libavcodec-dev libavformat-dev libswscale-dev libv4l-dev

1.1.4. sudo apt-get install libxvidcore-dev libx264-dev

1.1.5. sudo apt-get install libgtk-3-dev

*1.1.6. sudo apt-get install libcanberra-gtk**

1.1.7. sudo apt-get install libatlas-base-dev gfortran

2. Εγκατάσταση των headers της Python 3

2.1.1. sudo apt-get install python3-dev

3. Λήψη και εγκατάσταση της opencv και opencv_contrib (επιπλέον αλγόριθμοι που δεν περιλαμβάνονται στο πρότυπο περιβάλλον της opencv)

3.1.1. cd ~

3.1.2. wget -O opencv.zip https://github.com/opencv/opencv/archive/4.0.0.zip

3.1.3. wget -O opencv_contrib.zip

https://github.com/opencv/opencv_contrib/archive/4.0.0.zip

3.1.4. unzip opencv.zip

- 3.1.5. *unzip opencv_contrib.zip*
- 3.1.6. *mv opencv-4.0.0 opencv*
- 3.1.7. *mv opencv_contrib-4.0.0 opencv_contrib*
- 4. Εγκατάσταση του pip για τη διαχείριση πακέτων της Python 3
 - 4.1.1. *wget <https://bootstrap.pypa.io/get-pip.py>*
 - 4.1.2. *sudo python3 get-pip.py*
- 5. Εγκατάσταση περιβαλλόντων προγραμματισμού της Python 3
 - 5.1.1. *sudo pip install virtualenv virtualenvwrapper*
 - 5.1.2. *sudo rm -rf ~/get-pip.py ~/.cache/pip*
 - 5.1.3. *echo -e "\n# virtualenv and virtualenvwrapper" >> ~/.profile*
 - 5.1.4. *echo "export WORKON_HOME=\$HOME/.virtualenvs" >> ~/.profile*
 - 5.1.5. *echo "export VIRTUALENVWRAPPER_PYTHON=/usr/bin/python3" >> ~/.profile*
 - 5.1.6. *echo "source /usr/local/bin/virtualenvwrapper.sh" >> ~/.profile*
 - 5.1.7. *source ~/.profile*
- 6. Δημιουργία περιβάλλοντος για την εκτέλεση κώδικα της OpenCV και ενεργοποίησή του
 - 6.1.1. *mkvirtualenv cv -p python3*
 - 6.1.2. *workon cv*
- 7. Εγκατάσταση της βιβλιοθήκης πινάκων NumPy
 - 7.1.1. *pip install numpy*
- 8. Προετοιμασία εκτέλεσης του πηγαίου κώδικα της OpenCV
 - 8.1.1. *cd ~/opencv*
 - 8.1.2. *mkdir build*
 - 8.1.3. *cd build*
 - 8.1.4. *cmake -D CMAKE_BUILD_TYPE=RELEASE *
*-D CMAKE_INSTALL_PREFIX=/usr/local *
*-D OPENCV_EXTRA_MODULES_PATH=~/opencv_contrib/modules *
*-D ENABLE_NEON=ON *
*-D ENABLE_VFPV3=ON *
*-D BUILD_TESTS=OFF *
*-D OPENCV_ENABLE_NONFREE=ON *
*-D INSTALL_PYTHON_EXAMPLES=OFF *
-D BUILD_EXAMPLES=OFF ..
- 9. Αύξηση του SWAP για τη χρήση τεσσάρων πυρήνων του επεξεργαστή
 - 9.1.1. *sudo nano /etc/dphys-swapfile*

9.1.2. Αλλαγή του CONF_SWAPSIZE σε 2048

10. Επανεκκίνηση της λειτουργίας SWAP

10.1. `sudo /etc/init.d/dphys-swapfile stop`

10.2. `sudo /etc/init.d/dphys-swapfile start`

11. Εκτέλεση του πηγαίου κώδικα της OpenCV

11.1. `make -j4`

11.2. `sudo make install`

11.3. `sudo ldconfig`

11.4. Αλλαγή του CONF_SWAPSIZE σε 100MB και επανεκκίνηση της λειτουργίας SWAP όπως στο 9 και 10

12. Σύνδεση της OpenCV με το περιβάλλον της Python 3

12.1. `cd ~/.virtualenvs/cv/lib/python3.5/site-packages/`

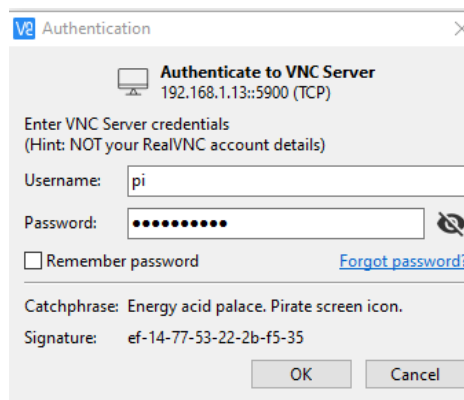
12.2. `ln -s /usr/local/python/cv2/python-3.5/cv2.cpython-35m-arm-linux-gnueabi.so cv2.so`

12.3. `cd ~`

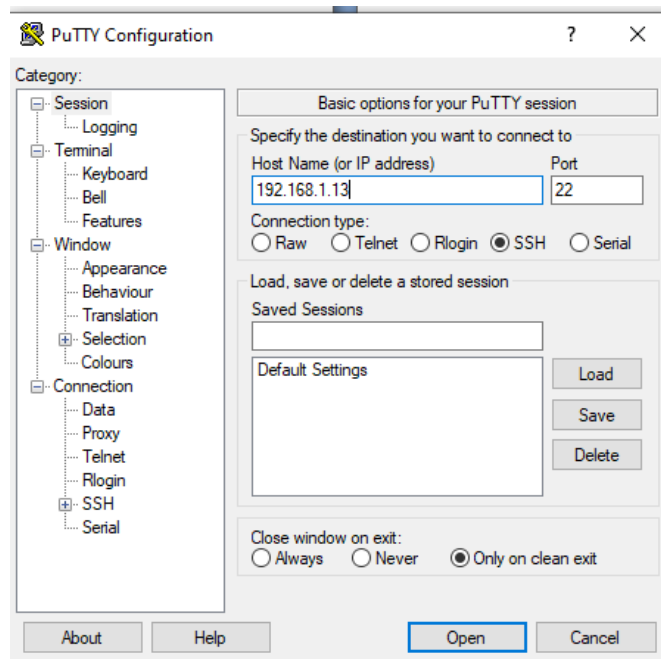
Για την είσοδο σε περιβάλλον εκτέλεσης κώδικα της OpenCV η εντολή που απαιτείται είναι **workon cv** ενώ για την έξοδο αντίστοιχα η εντολή **deactivate**.

B.3 Σύνδεση μέσω SSH και VNC

Δεδομένου ότι η χρήστης έχει ενεργοποιήσει τον VNC server και SSH server στο Raspberry Pi 4B όπως επίσης έχει εγκαταστήσει το RealVNC viewer και putty στην απομακρυσμένη συσκευή, η σύνδεση πραγματοποιείται όπως φαίνεται στις Εικόνα 19 και Εικόνα 20. Σε κάθε περίπτωση το όνομα χρήστη είναι "pi" ενώ ο κωδικός είναι αυτός που έχει χρησιμοποιηθεί για το λογαριασμό διαχειριστή του Raspberry Pi 4B.



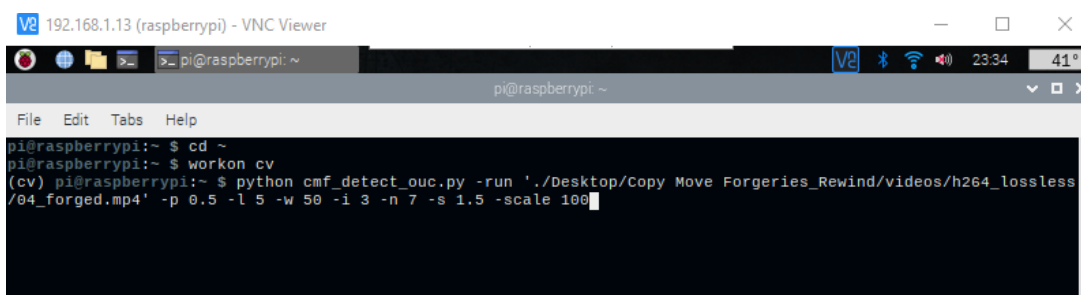
Εικόνα 19. Σύνδεση μέσω VNC viewer.



Εικόνα 20. Σύνδεση μέσω SSH/Putty.

B.4 Δημιουργία αρχείου pythοn και εκτέλεση

Για τη δημιουργία το αρχείου του εργαλείου της μεταπτυχιακής διατριβής ο χρήστης πρέπει να βρίσκεται στον αρχικό φάκελο (home folder) όπου θα δημιουργήσει ένα αρχείο κειμένου με όνομα **cmf_detect_ouc.py** στο οποίο θα αντιγράψει τον κώδικα από το Παράρτημα Α. Για την εκτέλεση του εργαλείου σε περιβάλλον της OpenCV πρέπει πρώτα να γίνει χρήση της εντολής **workon cv** και στη συνέχεια η εκτέλεσή του όπως έχει φαίνεται στην Εικόνα 21.



Εικόνα 21. Εντολές για την εκτέλεση του εργαλείου.

