



Τεχνολογικό
Πανεπιστήμιο
Κύπρου

Σχολή Μηχανικής και
Τεχνολογίας

Πτυχιακή εργασία

**Recurrent Convolutional Adversarial Networks for
Generative Modeling of Human Motion**

Λάμπρος Οδυσσέως

Λεμεσός, Μάιος 2017

ΤΕΧΝΟΛΟΓΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΗΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ Η/Υ
ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ

Πτυχιακή εργασία

Recurrent Convolutional Adversarial Networks for Generative
Modeling of Human Motion

του

Λάμπρου Οδυσσέως

Επιβλέπων Καθηγητής

Δρ. Σωτήριος Χατζής

Λεμεσός, Μάιος 2017

Πνευματικά δικαιώματα

Copyright © Λάμπρος Οδυσσέως, 2017

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της πτυχιακής εργασίας από το Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Η/Υ και Πληροφορικής του Τεχνολογικού Πανεπιστημίου Κύπρου δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Τμήματος.

Θα ήθελα να ευχαριστήσω ιδιαίτερα τον Δρ. Σωτήριο Π. Χατζή, για την άψογη συνεργασία καθώς και καθοδήγηση στα πλαίσια της διεκπεραίωσης αυτής της πτυχιακής εργασίας.

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια έχει αναδειχθεί η σημασία των Generative Adversarial Networks(GANs), καθώς και οι πιθανές χρήσεις τους στο εγγύτερο μέλλον. Σε αυτή την πτυχιακή εργασία πρόκειται να εξηγηθεί λεπτομερώς η ιδέα ενός μοντέλου μηχανικής μάθησης που βασίζεται κυρίως στα Generative Adversarial Networks αλλά καθώς και τα εργαλεία που χρησιμοποιήθηκαν για την ολοκλήρωση του.

Ένα Generative Adversarial Network είναι ένα μοντέλο μηχανικής μάθησης το οποίο αποτελείται από δύο βασικές συνιστώσες. Η πρώτη είναι το G(Generator) τμήμα και η δεύτερη είναι το D(Discriminator) τμήμα. Κατόπιν εκπαίδευσης, το μοντέλο είναι σε θέση να κάνει δύο πράγματα τα οποία ορίζονται από τις συνιστώσες του. Από την μία, μπορεί να παράγει δείγματα (samples) μέσω του G, τα οποία είναι σχετικά με τα δείγματα με τα οποία εκπαιδεύτηκε το μοντέλο, δηλαδή από τα δεδομένα εκπαίδευσης. Από την άλλη, είναι και σε θέση να διαχωρίζει αν ένα δείγμα προήλθε είτε από το G είτε από τα δεδομένα εκπαίδευσης, μέσω του D.

Το μοντέλο στο οποίο αναφέρεται αυτή η πτυχιακή εργασία συνδυάζει GANs μαζί με τεχνικές Convolution όσον αφορά την επεξεργασία των δεδομένων εισόδου για την εκπαίδευση. Επίσης, έχει γίνει ένας συνδυασμός με Recurrent Neural Networks(RNNs) και πιο συγκεκριμένα η εφαρμογή ενός Long Short Term Memory(LSTM) που επάγει την δυνατότητα δυναμικής ενθύμησης χρονοεξαρτημένων χαρακτηριστικών και εξαγωγής συσχέτισης μεταξύ ανθρώπινων κινήσεων.

Τέλος, το μοντέλο αυτό είναι σε θέση να παράγει δείγματα βίντεο που περιέχουν ανθρώπους να πραγματοποιούν κινήσεις, αλλά και να διαχωρίζει αν τα δείγματα αυτά προήλθαν από τον G ή από τα δεδομένα εκπαίδευσης.

Λέξεις κλειδιά: GANs, RNNs, LSTM, Convolution, Generator, Discriminator.

ABSTRACT

In recent years, the importance of Generative Adversarial Networks (GANs) has been discovered and their potential uses in the near future became obvious. In this thesis, the idea of a machine learning model based mainly on Generative Adversarial Networks will be explained in detail along with the tools used to complete it

A Generative Adversarial Network is a machine learning model that consists of two basic parameters. The first parameter is G (Generator) and the second is D (Discriminator). After training, the model is capable of doing two things that are defined by its parameters. First, it can produce samples through G, which are related to the samples which the model was trained with, the training data. On the other hand, it is also able to discriminate if a sample was produced either from G or from the training data through D.

The model referred to in this dissertation combines GANs along with Convolution techniques for processing input data for training. In addition, a combination of Recurrent Neural Networks (RNNs) and more specifically the fitting of a Long Short Term Memory (LSTM) has been made for a better convergence of the algorithm due to its ability to memorize features and extract dependencies between human movements.

Finally, this model is able to produce video samples that present human motion, but also to separate whether if these samples came from G or from the training data.

Keywords: GANs, RNNs, LSTM, Convolution, Generator, Discriminator.