UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Faculty of Science – Master's IP Assessment Form

Return this form with your Intention to submit form to:
Science Faculty Office
Level 6
PD Hahn Building

Contact:
Shahieda Samsodien
Tel: (021) 650 2713
E-mail: shahieda.samsodien@uct.ac.za

Note: This form must be returned to the Science Faculty Office with the Intention to Submit form.

In terms of the UCT Intellectual Property Policy and the Intellectual Property Rights from Publicly Financed R&D Act (No. 51 of 2008), it is a requirement to screen research outputs prior to public disclosure to ensure that any relevant intellectual property has been adequately protected prior to disclosure.

Additional information on the Act, the IP Policy and patentability can be obtained from the RCIPS website, see: www.rcips.uct.ac.za.

Once completed, this form will be reviewed by Research Contracts and Intellectual Property Services (RCIPS), who fulfil the role of the technology transfer office mandated by the Act.

Take note of thesis copyright issues on the RCIPS website (see: *Intellectual Property → IP for postgrad students* or http://www.rcips.uct.ac.za/rcips/ip/postgradsip).

Where necessary, it is possible for students to apply to the Faculty for confidentiality of a thesis to be maintained temporarily – without affecting graduation. This is typically associated with a need for patent protection. More details are available from RCIPS or the RCIPS website. As examiners need to be placed under a non-disclosure agreement this needs to be done well in advance of the submission of the dissertation.

Name	Christiaan Gerhardus Viljoen
Student number	VLJCHR004
E-mail address	Christiaan.viljoen@cern.ch; christiaan.viljoen.datasci@gmail.com
Telephone number	0847008414
Department and Faculty	Physics, Science
Dissertation title	Deep Generative Models for High Energy Physics Event Simulations and Convolutional Neural Networks for Particle Identification
Target hand-in date	30/11/2019
Supervisor/s name/s	Dr Thomas Dietel

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Abstract (may be copied from thesis)	This Masters project was focused on the application of deep learning techniques towards specific aspects of particle physics. Its two main aims: *particle identification* and *high energy physics event simulations* are pertinent to research avenues pursued by physicists working with the ALICE[1] TRD[2] detector, within the LHC[3] at CERN[4]. **Aims** More formally, the aims of this project were as follows: For particle identification: various neural networks were trained and assessed, to determine their ability to discriminate between electrons and pions, produced during proton-Lead (pPb) collisions conducted at the LHC in 2016, based on ADC[5] signal data produced as these particles were detected by the ALICE TRD. (Note that this work was done on uncalibrated raw TRD digits). For high energy physics event simulations: Geant4, a Monte Carlo toolkit used to simulate particle interactions with matter, was assessed in terms of how closely the simulated data it produces resembles true data taken by the TRD during collision events. In addition, as a step towards fast simulation, various deep generative modeling strategies were employed to produce simulated data samples which are likely under the observed (true) TRD data distribution. To this end, the following classes of latent variable models were prototyped: Generative Adversarial Networks, Variational Autoencoders and Adversarial Autoencoders. Data produced during these deep generative simulations were compared to real data in the same manner as that done for Geant4 data, in order to assess the feasibility of incorporating these types of models into future high energy physics event simulation software. **Summary of Results** Particle identification performance was defined by the ability of each neural network to minimize pion efficiency (ε_π, false positive rate), whilst maximizing electron efficiency (ε_e, true positive rate). A lower bound for the critical region (t_{cut}) in the distribution of $P(elec)$ predictions made by each neural network which results in $\varepsilon_e \approx 90\%$ was defined, in order to determine the ε_π for that neural network. The best set of results obtained, per momentum bin, was as follows: $\varepsilon_\pi = 1.2\%$ in the $P \leq 2\ GeV$ range; $\varepsilon_\pi = 1.14\%$ in the $2\ GeV < P \leq 3\ GeV$; and $\varepsilon_\pi = 1.51\%$ in the $3\ GeV < P \leq 4\ GeV$ range. In terms of results obtained for high energy physics event simulations, distinguishing Geant4 data from real data was a trivial task when compared to the task of particle identification. Similarly, data produced by deep generative models were easily distinguishable from real data; but the obtained results (especially for adversarial autoencoders) appear to be promising enough to pursue in future research. **Keywords** Deep Learning, Convolutional Neural Networks, Particle Identification, High Energy Physics Event Simulations, Generative Adversarial Networks, Variational Autoencoders, Adversarial Autoencoders, Geant4, ROOT, AliROOT, Tensorflow, Keras

Funder Who funded the project? Ignore any bursaries or scholarships received.	Personal funding
Patenting opportunities Describe any IP that you feel may be worth protecting.	None
Public disclosure Has there already been any presentation or publication of the work?	No

---

[1] A Large Ion Collider Experiment

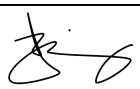
[2] Transition Radiation Detector

[3] Large Hadron Collider

[4] European Organization for Nuclear Research

[5] Analog to Digital Converter

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Research materials Are there any research materials (cultures, software, etc. – for full description, see UCT IP Policy) that have been developed? If so, broadly describe them and indicate who will become the custodian of them, if not your supervisor.	The following repositories contain code developed for my thesis: https://github.com/PsycheShaman/MSc-thesis https://github.com/PsycheShaman/deep-gen https://github.com/PsycheShaman/hpc-mini https://github.com/PsycheShaman/msc-hpc https://gitlab.cern.ch/cviljoen/msc-thesis-data (not public) https://github.com/PsycheShaman/trdpid https://github.com/PsycheShaman/Keras-GAN https://github.com/PsycheShaman/trdML-gerhard		
Signature		Date	27/10/2019

RCIPS Assessment - Office use only	
Protectable IP present?	
Public disclosure status	
Follow-up actions	
IP already protected?	
Funder issues?	