



**INFORMATICS  
INSTITUTE OF  
TECHNOLOGY**

INFORMATICS INSTITUTE OF TECHNOLOGY  
In Collaboration with  
UNIVERSITY OF WESTMINSTER

**Voice enhancement for voice artists using CycleGAN based  
non-parallel voice conversion**

A Dissertation by  
Pasindu Ekanayake  
W1742110 / 2018405

Supervised by  
Ms. Niwarthana Kariyabaduge

Submitted in partial fulfillment of the requirements for the BEng Software  
Engineering degree at the University of Westminster.

**May 2022**

## **Abstract**

Voice acting has long been regarded in the entertainment industry as an important section, and it is vital to the success of any product. Inexperienced and sometimes even experienced casting directors struggle to identify the best voice actor for the job, and indie producers find it costly to cast many outstanding voice actors at once. As a result, a voice enhancement technology that can change a given voice into a target voice would allow casting directors to focus on finding skilled voice actors rather than obsessing over the actor's voice. Also, it would help producers save money by opening the option for them to cast a small number of competent performers and change their voices for each role.

This research involves incorporating voice conversion techniques to address the drawbacks identified in the voice casting process. Voice conversion is a technique that retains linguistic information like words, pronunciation, and accent while transforming non-linguistic information like voice quality. The research will conduct experiments to determine the model's effectiveness on Sinhala language because voice conversion technology has yet to be introduced to the Sinhala language.

This study will provide a system that can enhance the voice of a voice artist using state-of-the-art voice conversion models. This will examine how the field of voice conversion can be applied to this problem domain, as well as the most effective voice conversion approaches and a review of existing models. Casting directors and producers will benefit from this system because it will save casting time and expenses.

**Keywords**— Voice Casting, Voice Acting, Generative Adversarial Networks, Voice Conversion, Source to Target, Generative Model