

# Overcoming Challenges with Automatic Video Colourisation



*Three black and white frames (top) taken from “His Mother” (1912) and their corresponding colourised versions (bottom).*

## Introduction

Colourisation is the process of adding colour to black and white media. Its origins date back to the early 20th century for movies where they painted directly onto film stock with brushes. In recent times it has gained more popularity, possibly due to the modern audiences' desire for a colourful experience. Nationally, we have become accustomed to media in colour here since the first colour broadcasts in the late 1960s, and the first outside broadcast (the Eurovision Song Contest) by RTÉ in 1972.

From social media to cinema, books to TV series, it is hard to find a medium that is not touched by some form of colourisation. On the social media side, there are Twitter, Instagram and Facebook accounts specialising in posting colourised material (e.g., Vid2Color Research). There are multiple bestselling books including the “Old Ireland in Colour” series publishing colourised work. Regarding art, galleries such as “The Klimt Colour Enigma” are showcasing their colourised work (of black and white photographs of extant paintings). On television, there are documentaries such as “World War II in Colour” and the Irish “Revolution in Colour”. And finally, wherein lies the main focus of our work, there is the

film industry, where large-scale movies and shows such as “Sherlock Holmes” and “Doctor Who” are also being enhanced through colourisation.

## What are the challenges with colourisation?

Unfortunately, colourisation is not a simple process. It has been largely manual and time consuming. It is meticulous work, which requires a keen eye for detail. From a technical perspective, it is an ill-posed problem. There are multiple plausible colourisations for any given black and white picture.

Let us take an example. Imagine you see a black and white image of a car: how will you know what colour it is? The answer is: you cannot without further information. Perhaps someone might choose silver or black, because a bright yellow car would be incongruent in most settings. However, for something like a military uniform, it is important that the colours are historically accurate, or else the video takes on a whole different meaning.

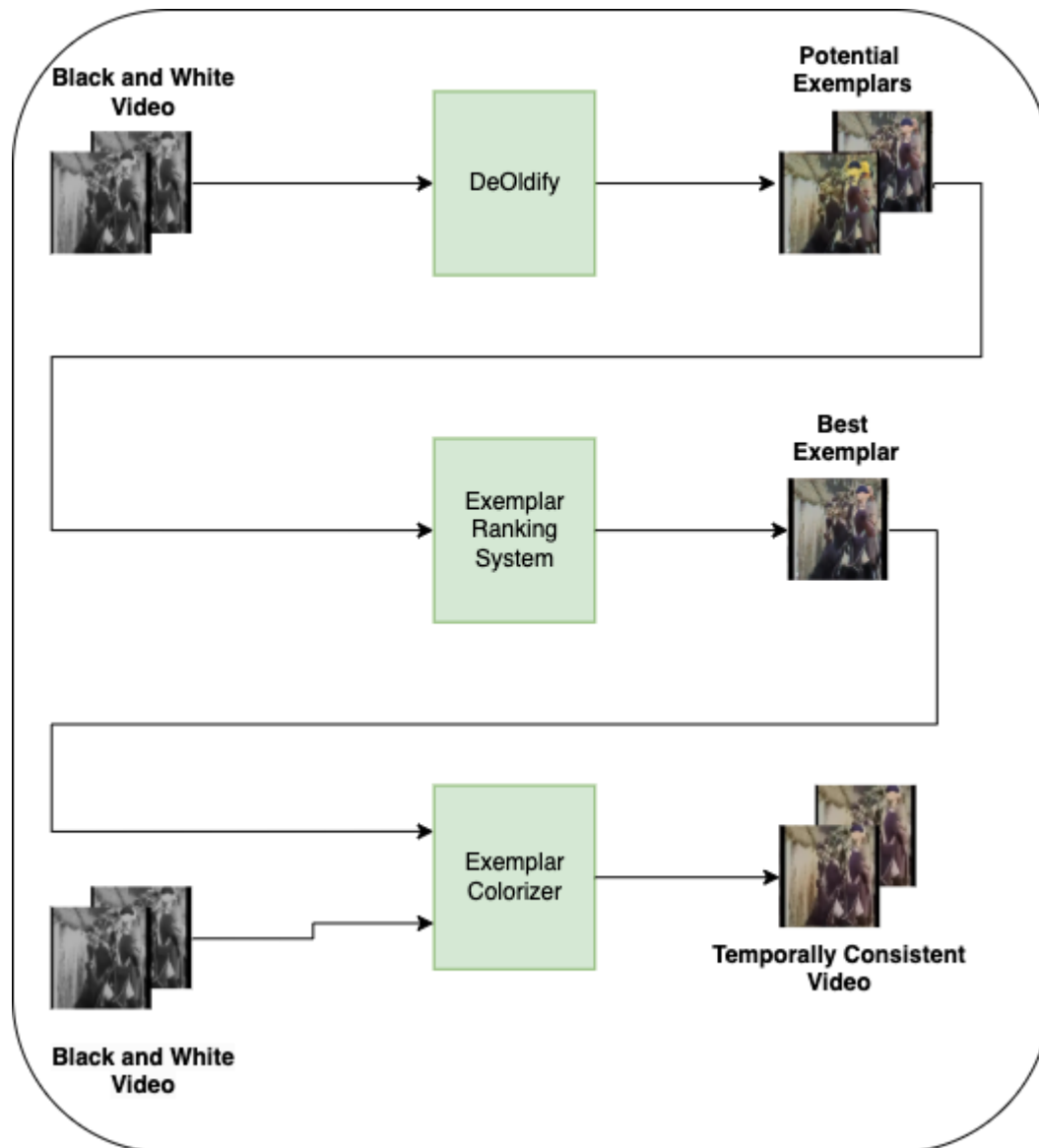
Aiming to tackle these challenges, automatic colourisation with context has been proposed. Automatic colourisation attempts to reinvent the colourisation process in a more efficient, less labour-intensive way. Initially it was designed for images, but it has also been extended to videos. The performance it achieves in a comparatively short period of time has the potential to revolutionise this industry.

One issue automatic colourisation faces is consistency through its media, whether that be an image or a video. It demands that the model must have a good understanding of the world, for example, it requires the model to realise that a person’s hand is a part of the person and should therefore be coloured with the same skin colour as their face. It requires the model to understand the idea of persistence. For example, one must understand that objects (e.g., clothes) in a video sequence should not change colour from frame to frame as this results in a video that has an artificial looking flicker effect.

## Current methods

There are three main automatic colourisation methods. Each method is inspired by how humans colourise. These methods are scribble-based, exemplar-based and learning-based: *Scribble-based methods* work by the user inputting colour hints into the system. These colour hints then guide the colourisation process. Photoshop has such a method built into their software whereby you click on points in an image, and it spreads that colour to what it determines are the appropriate regions. *Exemplar-based methods* are best explained by an example. When employing this method on colourising an image of a black and white tree, the system is given a colour image of a tree towards which it can guide its colourisation. *Learning-based methods* use a large data set of images or video on which they base their colourisations on. It does this by updating the weights of its neural network model in a deep learning process. Each method, however, has its own limitations. Scribble- and exemplar-based methods require human input. However, learning-based methods rely on their data sets and do not consider human feedback, short of retraining the model on specific data sets.

# Ideas for improvements



*Network diagram of our framework*

The particular problem that we attempted to solve was lack of consistency in a video sequence. For a more detailed explanation, see our paper “Towards Temporal Consistency in Automatic Video Colourisation”. To summarise, we created a hybrid system of exemplar- and learning-based systems. We employed a learning-based system inspired by DeOldify to create a suitable exemplar, then used a “best” exemplar-based system to ensure temporal consistency. We developed a novel exemplar selection algorithm. The aim of all this was to reduce flicker in automatic video colourisation and therefore create more consistent colourisations.

## But, what's a good or a bad colourisation result?

Like all novel solutions, it was not smooth sailing from start to finish. We did encounter a few difficulties, with the main one being, how to measure colourisation performance. The human brain is complex, and it is difficult to evaluate colour as a number. It is quite an arbitrary notion, the idea of a “good colourisation” versus a “bad colourisation” and is often different depending on who you ask. Of course, this element of subjectivity does not conform to an engineering project where we focus on results and performance. Therefore, we decided to use reference and non-reference metrics. Reference metrics calculate the difference from ground truth. Although useful, they cannot be applied to most colourisation projects as often there is no ground truth (version in colour) available. The non-reference metrics are image quality analysis. These families of metrics, including NIQE (Natural Image Quality Evaluator) and BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator), are useful as they give a general idea of colourisation consistency and authenticity.

## Our experiments and some results

The black and white video is initially passed through an open-source video colourisation network named DeOldify. This produces potential exemplars. This exemplar frame is chosen in the Exemplar Ranking System which selects the “best” exemplar based on its image quality analysis techniques. This “best” exemplar is then passed through the exemplar-based colouriser, along with the original black and white video, to produce a more temporally consistent coloured video.

Method	NIQE ↓	BRISQUE ↓
DeOldify	16.23	65.43
Min_SEA	16.12	66.31
Avg_SEA	16.15	66.32
Max_SEA	16.16	66.33
BRISQUEd	16.14	66.33
NIQEd	16.15	66.33
Ours	<b>15.01</b>	<b>60.02</b>

*Quantitative non-referenced comparisons on old movies (from first decades of 20th century)*

The end result was an 8.5% increase in non-referenced image quality over the previous state of the art. We published this work in the Irish Machine Vision and Image Processing Conference (IMVIP), where we received the “Best Presentation Award”.

## What we learned

This project gave us a great insight into the compute intensiveness of testing and training deep learning-based systems. We learned that there is a detailed development process involved in training large neural networks. We also have huge admiration for the DeOldify project and team, who have pioneered work in automatic video colourisation. They have contributed greatly to our interest in colourisation, and we strive to develop their work further.

Going forward we plan to continue to develop this project by conducting a large-scale study to determine the performance of the system on human evaluation. Additionally, we plan to develop more capabilities into the system such as optical flow-based methods.

## Biographies

**Rory Ward** is a second year PhD researcher in the Visual Media Enhancement Laboratory, University of Galway, where he is supervised by Prof. John Breslin, and funded by the SFI Centre for Research Training in Artificial Intelligence. His research explores novel techniques for efficient and accurate colourisation of black and white media. Rory holds a BE in Electronic and Computer Engineering from the University of Galway.

**John Breslin** is a Professor of Electronic Engineering at the University of Galway and Co-PI and the Insight and Confirm SFI Centres. He has co-authored over 290 publications, and co-created the SIOC ontology, implemented in hundreds of applications on thousands of websites with millions of data instances. He is co-founder of boards.ie, adverts.ie, and the PorterShed (Galway City Innovation District CLG). He is co-author of Old Ireland in Colour.