

## Attention-based Style Transfer for Korean Traditional Painting

Jin-Woo Cha\*, Kyu-Cheol Cho\*\*

\*Student, Dept. of Computer Science, Inha Technical College, Incheon, Korea

\*\*Professor, Dept. of Computer Science, Inha Technical College, Incheon, Korea

## [Abstract]

Neural Style Transfer is being utilized in various fields while preserving the form of the image as deep learning network that changes aesthetic elements such as color and texture. Moreover many Neural Style Transfer studies have been conducted based on famous western art datasets and has shown excellent performance in westernization transfer and has contributed greatly to the development of influence in that cultural sphere. But transferring traditional korean painting styles such as ink-and-wash painting is difficult through existing methods, and related research is lacking. This paper proposes a novel network that effectively reflects the characteristics and margins of light and shade of images with combining CycleGAN and Self-Attention mechanisms. We also developed a web service that can generate images having the characteristics of korean painting using the proposed network. We hope that this service will encourage anyone to take an interest in this research topic.

▶ **Key words:** Neural Style Transfer, Self-Attention, Traditional Korean Painting, CycleGAN, Deep Learning

## [요 약]

Neural Style Transfer는 이미지의 형태를 보존하면서 색채와 질감 등의 미학적 요소를 변경하는 딥러닝 네트워크로 다양한 분야에서 활용되고 있다. 대부분의 Neural Style Transfer는 유명 서양 미술 데이터 세트를 기반으로 연구되어 왔으며 서양화 전이에 탁월한 성능을 보여 해당 문화권 영향력 발전에 크게 기여하고 있다. 그에 반해 수묵화 등 한국 전통 회화 양식 전이는 기존 네트워크로 수행하기 어려우며 관련 연구가 부족하다. 본 논문은 CycleGAN과 Self-Attention 메커니즘을 결합하여 농담의 특성과 여백을 효과적으로 반영하는 새로운 네트워크를 제안한다. 또한 제안된 네트워크를 사용하여 이미지를 생성할 수 있는 웹 서비스를 개발함으로써, 누구나 본 연구 주제에 관한 관심을 가질 수 있도록 하였다.

▶ **주제어:** 신경망 스타일 변환, Self-Attention, 한국 전통 회화, CycleGAN, 딥러닝

- 
- First Author: Jin-Woo Cha, Corresponding Author: Kyu-Cheol Cho
  - \*Jin-Woo Cha (chajinwoo.chajinwoo@gmail.com), Dept. of Computer Science, Inha Technical College
  - \*\*Kyu-Cheol Cho (kccho@inhac.ac.kr), Dept. of Computer Science, Inha Technical College
  - Received: 2025. 03. 26, Revised: 2025. 05. 02, Accepted: 2025. 08. 22.

## I. Introduction

Neural Style Transfer(NST)는 이미지의 구조를 보존하면서 목표 특성을 전이하는 영상 처리 기법이다. 2016년 Gatys 등[1]이 처음 제안한 이후 실시간 처리와 고해상도 이미지 생성 등의 발전을 거듭해 왔다. 특히 이미지 특징 추출 기술의 발전으로 콘텐츠와 스타일의 독립적 조작이 정교해졌고, 이는 디지털 아트와 미디어 콘텐츠 제작의 새로운 패러다임을 제시했다.

일반 이미지를 서양 회화 양식으로 전이하는 기술은 이미 높은 성과를 거두어 예술의 새로운 표현 방법으로 자리잡았다[2]. 한국 전통 회화의 경우에도 NST 기술을 활용하여 문화적 가치를 더할 수 있도록 다양한 연구가 필요한 상황이다.

그러나 대부분의 NST 연구는 서양화 스타일에 최적화된 네트워크를 설계하는 방향으로 발전해 왔다[3]. 한국화를 포함한 동양화 관련 연구는 상대적으로 제한적이며, 이는 한정된 데이터 세트와 딥러닝 모델로 특징을 포착하기 어려운 동양화 특유의 화풍에서 기인한다.

본 연구에서는 한국화의 대표적인 회화 양식인 수묵화에 초점을 맞춘다. 수묵화는 붓의 질감이 직관적으로 표현되며 적극적으로 여백을 활용한다는 점에서 서양화와의 차이점이 뚜렷하기에 본 연구에서 다루어야 할 기술적 과제를 명확히 제시한다.

수묵화와 서양화의 예술적 특성은 획과 여백으로 대표할 수 있다. 획의 차이는 먹과 붓의 성질이 서양화에서 쓰이는 도구와 다르기 때문에 나타난다. 한 번의 붓질 안에서 다양한 진하기와 굵기, 화선지와 의 상호작용으로 만들어지는 번짐 효과가 나타난다. 이러한 선은 그 자체로 깊이감과 생동감을 만들어낸다[4][5].

여백은 동양 회화를 대표하는 특징적 요소이다. 의미론적 관점에서는 그려진 부분의 의미를 확장한다. 구조적 관점에서는 여유와 휴식의 공간으로서 그려진 대상을 부각하고 눈의 피로감을 줄인다. 철학적으로는 음양의 조화와 자연스러움을 추구하여 동양 철학의 사상을 반영한다. 즉, 여백은 그려지지 않은 공간이 아닌 작품 일부로서 중요한 의미가 있다.

이러한 특징들은 기존의 Neural Style Transfer 알고리즘으로는 표현하기 어려운 요소이다. 수묵화의 추상적 표현을 효과적으로 포착하지 못한다는 문제가 선행 연구들에서 지속적으로 지적되어 왔다[6][7]. 하나의 획에 포함된 미세한 강약의 변화는 소실되어 개별적인 특징으로 인식하지 못한다. 여백에 내포된 의미와 역할을 반영하지 못

하고 단순한 배경으로 인식하여 수묵화의 공간적 구성을 왜곡한다.

본 논문에서는 간소화된 CycleGAN에 Self-Attention 매커니즘을 도입하여 한국 수묵화의 특성을 효과적으로 포착하는 새로운 네트워크를 제안한다.

본 논문의 구성은 다음과 같다. 2절에서는 기존 NST 연구가 발전해 온 과정과 최신 연구 동향을 정리하고 수묵화 생성을 위한 데이터셋을 소개한다. 3절에서는 어텐션 매커니즘을 적용한 GAN 구조의 수묵화 NST 네트워크를 제안하며, 정량적 평가를 통해 수묵화 스타일 전이에 효과적임을 보여준다. 4절에서는 연구의 결과와 한계를 분석하고 향후 연구 방향에 대해 논의한다.

## II. Preliminaries

NST는 이미지의 사실적인 렌더링이 아닌 예술적 표현 스타일을 컴퓨터로 구현하는 Non-Photorealistic Rendering 분야 중 Image-Based Artistic Rendering (IB-AR)에 속한다. 본 논문은 Kyprianidis 등[8]이 정의한 분류 체계를 따른다.

### 2.1 Image optimization based Online Neural Methods

Gatys 등[1]은 CNN 특징 추출 모델을 활용하여 이미지의 스타일과 콘텐츠를 분리하는 새로운 네트워크를 제안했다. 사전 학습된 VGG[9] 네트워크의 각 층에서 입력 이미지의 특징을 계층적으로 추출하였는데, 상위층에서 콘텐츠의 특징을 추출하고, 여러 층에서 동시에 스타일 특징을 추출하였다. 콘텐츠의 고수준 정보와 스타일의 질감, 색상, 시각적 패턴 등을 동시에 포착하여 높은 품질의 스타일 변환 결과를 생성할 수 있었으며 NST 연구의 초석이 되었다.

이 기법은 콘텐츠 손실(content loss)과 스타일 손실(style loss)이라는 두 가지 주요 손실 함수를 최적화하는 방식으로 작동한다. 콘텐츠 손실은 생성된 이미지와 원본 콘텐츠 이미지 사이의 고수준 특징 차이를 측정하며, 스타일 손실은 Gram 행렬을 사용하여 스타일 이미지와 생성된 이미지 간의 텍스처 차이를 계산한다. 이러한 손실 함수들은 L-BFGS와 같은 최적화 알고리즘을 통해 반복적으로 최소화되며, 이 과정에서 백색 노이즈로 초기화된 이미지가 점진적으로 스타일이 적용된 결과물로 변환된다.

이 방법은 우수한 결과물을 생성할 수 있지만, 각 스타일 변환마다 수백 번의 반복 최적화 과정이 필요하여 계산

비용이 높고 실시간 처리가 어렵다는 단점이 있다. Li 등 [10]은 Markov Random Field(MRF)와 CNN 특징을 결합하여 보다 지역적인 스타일 특성을 보존하는 방법을 제안하였다.

이러한 이미지 최적화 기반 방법들은 높은 품질의 결과물을 생성할 수 있으나, 계산 효율성 측면에서의 한계로 인해 점차 모델 최적화 기반 방법들로 연구의 중심이 이동하게 되었다.

## 2.2 Model optimization based Online Neural Methods

Johnson 등[11]은 사전 학습된 변환 네트워크를 사용하여 최적화 과정을 단일 순전파로 대체하는 방법을 제안하여 모델 최적화 기반 방법의 중요한 기반을 마련했다. Huang 등[12]은 AdaIN(Adaptive Instance Normalization) 네트워크를 제안하였다. AdaIN은 콘텐츠 이미지의 특징 통계량과 스타일 이미지의 채널별 통계량을 사용하여 복잡한 반복 최적화 과정 없이 효율적인 스타일 변환을 수행한다. 실시간 처리가 가능하면서도 높은 품질의 이미지를 생성할 수 있어 주목받았으며, 다양한 스타일에 적용할 수 있는 유연성을 제공한다. Ulyanov 등[13]은 Instance Normalization을 도입하여 배치 정규화의 한계를 극복하고 스타일 변환의 품질을 향상시켰다. 훈련 단계에서 모델의 가중치를 최적화함으로써 추론 시 계산 비용을 크게 줄일 수 있었다.

모델 최적화 기반 방법은 계산 효율성과 안정적인 생성 능력으로 가장 오랫동안 활발한 연구가 진행되었으며, 다양한 응용 분야에서 중요한 성과를 이루어냈다. 특히 Dumoulin 등[14]의 다중 스타일 처리 기법과 Li 등[15]의 Universal Style Transfer 연구는 이 분야의 발전 가능성을 더욱 확장시켰다.

그러나 이러한 모델들은 주로 자연 이미지와 서양화 데이터셋으로 학습되어 수묵화의 추상적 특징, 특히 붓의 농담 변화나 여백의 의미론적 가치를 효과적으로 포착하기 어렵다는 한계가 있다. 농담의 미세한 변화는 CNN의 특징 추출 과정에서 단순화되거나 손실되기 쉽고, 여백의 의도적 활용은 단순한 배경으로 오인될 수 있다.

## 2.3 GAN based Methods

GAN(Generative Adversarial Network)[16]는 이미지 생성 분야에 혁신적인 변화를 가져왔다. 특히 Zhu 등[17]이 제안한 CycleGAN은 스타일 변환 전후의 이미지 쌍 데이터 없이 훈련이 가능함을 입증했다는 점에서 의미가 크다. 순환 일관성 손실(Cycle consistency loss)를 도입하

여 스타일 변환된 이미지가 다시 원본으로 복원될 수 있도록 일관성을 보장하고 안정성과 품질을 높였다.

GAN은 한국화 스타일 변환 연구에도 의미가 크다. 고민삼 등[18]은 GAN을 개선한 한국 수묵화 생성 모델인 GANMOOK을 제안하여 딥러닝 기반 스타일 변환 알고리즘을 한국화 양식에 적용할 수 있다는 가능성을 제시하였다. 특히 한정된 데이터셋을 사용했음에도 자연 이미지를 수묵-산수화로 변환하는데 뛰어난 성능을 보였다.

GAN 기반 스타일 변환 방식은 실사적인 이미지 생성 능력과 도메인 간 효과적인 매핑이 가능하다는 장점이 있으나, 모드 붕괴(Mode collapse) 현상과 훈련 불안정성이라는 단점도 존재한다. 수묵화와 같이 추상적 표현이 많은 도메인으로 변환할 때는 세부 표현이 과도하게 단순해지거나 불필요한 텍스처가 생성될 수 있다.

## 2.4 Stroke based Methods

스트로크 기반 방법론은 회화의 물리적 붓질을 시뮬레이션하여 디지털 이미지를 회화로 변환하는 접근법이다. 이는 실제 화가의 붓놀림을 모방함으로써 자연스러운 회화적 표현을 구현하고자 한다. 스트로크 기반 방법은 크게 규칙 기반(rule-based)과 학습 기반(learning-based) 접근법으로 나눌 수 있다. Li와 Wand[19]는 MRF(Markov Random Field)와 CNN을 결합하여 텍스처 합성 방식의 스트로크 생성 방법을 제안했다.

Zou 등[20]은 스트로크 매개변수를 순차적으로 예측하는 강화학습 기반 모델을 제안하였다. 이 방법은 스트로크 단위의 렌더링을 통해 실제 그림을 그리는 과정을 모방한다. 모델은 스트로크의 모양, 색상, 크기, 위치, 투명도 등의 매개변수를 예측하고, 이를 기반으로 스트로크를 캔버스에 순차적으로 적용한다. 이러한 방식으로 회화의 질감 표현을 극대화하고, 실제 화가의 작업 과정과 유사한 결과물을 생성할 수 있다.

그러나 이러한 스트로크 기반 방법론은 수묵화 표현에 있어 몇 가지 한계를 갖는다. 첫째, 비슷한 모양의 스트로크를 여러 겹 덧씌우는 방식을 사용한다. 이는 한 획의 농담 변화와 번짐 효과가 중요한 수묵화의 표현 방식과 상충한다. 둘째, 획의 형태와 농담 변화를 독립적으로 제어하기 어렵다. 셋째, 화선지와 먹의 상호작용에서 발생하는 미세한 번짐과 흡수 효과를 물리적으로 시뮬레이션하는데 제약이 있다.

## 2.5 Transformer based Methods

Transformer[21]는 2017년 Vaswani 등이 제안한 자기주의 메커니즘(Self-Attention mechanism)을 기반으로 한 신경망 아키텍처이다. Transformer는 RNN 기반 모델의 한계를 극복하고 병렬 처리에 최적화된 구조를 제공하여 학습 효율성과 성능을 크게 향상시켰다.

Dosovitskiy 등[22]은 ViT(Vision Transformer)를 제안하였다. 자연어 처리에 사용되던 Transformer를 컴퓨터 비전 분야로 확장하여 CNN 없이도 이미지 분류 작업에서 뛰어난 성능을 달성할 수 있음을 증명했다. ViT는 이미지를 패치로 분할하고 각 패치를 토큰으로 취급하여 Transformer 인코더에 입력함으로써 이미지의 전역적 맥락을 효과적으로 포착한다.

Transformer 기반 방법론의 가장 큰 장점은 자기주의 메커니즘을 통해 이미지의 각 부분 간의 장거리 관계를 효과적으로 모델링할 수 있다는 점이다. 이는 수묵화의 여백과 진한 부분 간의 상호 관계, 획의 연속성과 같은 전역적 특성을 포착하는 데 유리하다.

한국 전통 회화 스타일 변환에 있어 Transformer 기반 모델은 여백의 의미론적 중요성과 획의 전체적인 흐름을 파악하는 데 유리하지만, 연산비용이 높고 대량의 학습 데이터가 필요하다는 단점이 있다. 또한, 한국화 특유의 미세한 질감 표현과 붓의 움직임에 따른 시간적 요소를 완벽히 포착하기 위해서는 추가적인 구조적 개선이 필요하다.

## 2.6 Recent Research Trends

최근 연구들은 다양한 방향으로 발전하고 있다. Park 등[23]은 2022년 SWAGAN을 제안하여 sliding window attention을 통해 고해상도 이미지의 스타일 변환 성능을 크게 향상시켰다. 이 연구는 기존 attention 메커니즘의 계산 복잡도 문제를 해결하면서도 전역적 맥락을 효과적으로 포착할 수 있음을 보여주었다.

Chen 등[24]은 2023년 ArtFlow라는 새로운 접근법을 제시했다. 이들은 normalizing flow를 활용하여 스타일과 콘텐츠의 분리를 더욱 정교하게 수행할 수 있는 방법을 개발했다. 특히 동양화와 같은 추상적 표현에서도 우수한 성능을 보였다는 점에서 본 연구와 밀접한 관련이 있다.

Kim 등[25]은 2024년 StyleFormer를 발표하여 순수 Transformer 기반의 스타일 변환 모델을 제안했다. 이 연구는 기존 CNN 기반 방법들이 포착하지 못하는 장거리 의존성을 효과적으로 학습할 수 있음을 입증했으며, 특히 여백의 활용이 중요한 동양화 스타일 변환에서 혁신적인 결과를 보여주었다.

Wang 등[26]은 2023년 비대칭 Cycle-consistent GAN을 이용한 수묵화 스타일 변환 방법을 제안했다. 이 연구는 실사 사진과 수묵화 도메인 간의 정보 차이를 고려하여 기존 대칭적 구조의 한계를 극복하고자 했다. 특히 붓터치 스타일, 먹의 확산, 세부 보존 측면에서 우수한 성능을 보였다.

Fu 등[27]은 2024년 attention 메커니즘과 GAN을 통합한 이미지 스타일 변환 모델을 제안했다. 이 연구는 cycle-consistent 네트워크와 attention 메커니즘을 결합하여 서로 다른 스타일과 콘텐츠 간의 장거리 의존성 문제를 해결하고자 했다.

또한 동양화 스타일 변환 분야에서는 전통적 특징 보존과 현대적 기법의 융합이 주요 과제로 대두되고 있다. 특히 saliency-based attention 메커니즘은 전통 회화의 핵심 요소를 선별적으로 강조할 수 있어 문화적 정체성 보존에 중요한 역할을 하고 있다[28].

본 연구는 이러한 최신 연구들과 달리 한국 수묵화의 고유한 특성인 농담과 여백에 특화된 Self-Attention 메커니즘을 CycleGAN과 결합한다는 점에서 차별성을 갖는다. 또한 간소화된 구조를 통해 계산 효율성을 높이면서도 문화적 특성을 효과적으로 보존할 수 있는 방법을 제시한다.

Transformer 기반 방법론의 가장 큰 장점은 자기주의 메커니즘을 통해 이미지의 각 부분 간의 장거리 관계를 효과적으로 모델링할 수 있다는 점이다. 이는 수묵화의 여백과 진한 부분 간의 상호 관계, 획의 연속성과 같은 전역적 특성을 포착하는 데 유리하다.

한국 전통 회화 스타일 변환에 있어 Transformer 기반 모델은 여백의 의미론적 중요성과 획의 전체적인 흐름을 파악하는 데 유리하지만, 연산비용이 높고 대량의 학습 데이터가 필요하다는 단점이 있다. 또한, 한국화 특유의 미세한 질감 표현과 붓의 움직임에 따른 시간적 요소를 완벽히 포착하기 위해서는 추가적인 구조적 개선이 필요하다.

## 2.7 Korean Traditional Ink-Painting Dataset

경희대학교에서 주관하여 2022년 AIHUB에 공개된 한국 전통 수묵화 화풍별 제작 데이터(이하, 수묵화 데이터)는 풍경 사진, 애니메이션 장면, 유명 서양화 등을 화풍별 수묵화로 재현하여 한국화의 디지털화에 크게 기여했다. 이 데이터셋은 약 7,000장의 이미지와 200건의 동영상 데이터가 표현 기법, 종이 유형 등이 구체적으로 기술된 어노테이션과 함께 제공되어 향후 다양한 연구에 사용될 수 있는 발판이 되었다.

### III. The Proposed Scheme

#### 3.1 System Architecture

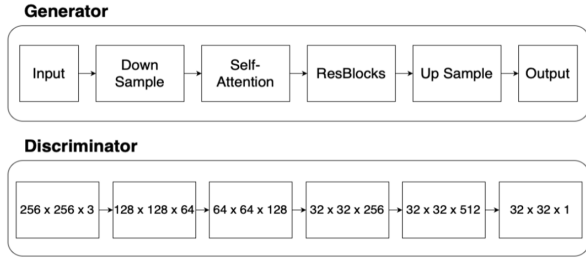


Fig. 1. System Architecture

본 연구는 간소화된 CycleGAN에 Self-Attention 메커니즘을 도입하여 한국 수묵화의 특성을 효과적으로 포착할 수 있는 새로운 네트워크를 제안한다. [그림 1]은 본 모델의 Generator와 Discriminator의 구조를 보여준다. 수묵화에서 여백은 단순한 빈 공간이 아닌 의미를 갖는 구성 요소이다. Self-Attention의 어텐션 가중치  $\alpha_{ij}$ 는 다음과 같이 계산된다.

$$\alpha_{ij} = \text{softmax}(Q_i \cdot K_j^T / \sqrt{d_k})$$

이때 여백 영역의 Query가 그려진 영역의 Key와 상호 작용하여 전체적인 조화를 파악할 수 있다. 기존 CNN의 지역적 수용장과 달리, 어텐션은 이미지 전체 영역 간의 관계를 직접 모델링한다. 한 획 내에서 발생하는 미세한 농담 변화는 연속적인 특성을 갖는다. Self-Attention의 Value 벡터  $V$ 는 이러한 연속성을 다음과 같이 보존한다.

$$\text{Output}_i = \sum_j \alpha_{ij} \cdot V_j$$

여기서  $\alpha_{ij}$ 가 공간적으로 인접한 픽셀들에 대해 높은 가중치를 부여하면서도, 전체적인 농담의 흐름을 유지할 수 있다.

#### 3.2 Generator

본 연구에서는 CycleGAN의 양방향 변환 모듈을 단방향 변환으로 간소화하고 업샘플링 레이어를 축소하여 모델의 복잡성을 줄이고 학습 효율성을 높이고자 하였다. 기존 CycleGAN 대비 파라미터는 11.4M에서 6.8M으로 40.4% 감소했으며 37.8% 감소한 메모리 사용량으로 학습 가능하다. 이로 인해 응용 서비스와 에이전트 개발에 적용하기 용이하다.

Self-Attention 모듈은 다운샘플링 레이어와 ResBlock 사이에 배치되어 있다. 이는 특징맵의 모든 위치 간의 관계를 계산하여 전역적 문맥 정보를 포착한다. Query,

Key, Value 투영을 통해 attention map을 생성하며, 이는 수묵화의 농담 표현과 여백의 조화를 잡는 중요한 역할을 한다.

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T)V$$

$Q$ 는 특징맵에서 주목할 부분을 결정하는 질의 벡터이다.  $K$ 는 모든 위치의 특징들을 표현하는 키 벡터로,  $Q$ 와 유사도 계산에 사용된다.  $V$ 는  $K$ 와 짝을 이루는 실제 특징 값이다.  $Q$ 와  $K$ 의 유사도에 따른 가중치가 부여되며 최종 출력에 반영된다.  $T$ 는 행렬의 전치 연산이다.

네 개의 ResBlock은 Skip-Connection을 통해 그래디언트 소실 문제를 완화하고 원본 이미지의 구조적 정보를 보존한다. 각 ResBlock은 두 개의  $3 \times 3$  합성곱 레이어로 구성되며, Instance Normalization과 ReLU를 포함한다.

업샘플링 과정에서는 전치 합성곱(Transposed Convolution)을 사용하여 특징맵의 해상도를 점진적으로 복원한다. 마지막으로  $7 \times 7$  합성곱 레이어와 tanh 활성화 함수를 통해 최종 결과물을 생성한다.

#### 3.3 Discriminator

Discriminator는  $70 \times 70$  PatchGAN[29] 구조를 채택하였다. 이는 이미지를 패치 단위로 판별하여 지역적인 스타일의 일관성을 평가한다. 다섯 개의 합성곱 레이어로 구성되며, 첫 번째 레이어를 제외한 모든 레이어에 Instance Normalization을 적용하였다. LeakyReLU 활성화 함수 ( $\alpha=0.2$ )를 사용하여 그래디언트 소실을 방지하고 학습의 안정성을 높였다.

#### 3.4 Loss function

본 모델은 MSE 기반의 GAN Loss를 사용하여 Generator와 Discriminator를 동시에 학습한다. Discriminator는 Generator가 생성한 이미지와 실제 이미지를 구분하도록, Generator는 Discriminator가 구분하지 못하는 정교한 이미지를 생성하도록 적대적으로 학습된다. 손실 함수의 수식은 아래와 같다.

$$L_{MSE} = \text{MSE}(D(x), \text{target})$$

$x$ 는 Discriminator에 입력되는 이미지이다. target은 실제 이미지의 경우 1, 생성된 이미지의 경우 0으로 정의된다. 최종적으로 MSE 함수를 사용하여 실제 이미지와 생성된 이미지의 차이를 수치화한다.

### 3.5 Training Details

데이터셋은 CycleGAN의 monet2photo[5] 세트 중 자연 이미지 세트를 콘텐츠 이미지로, 한국 수목화 데이터셋을 스타일 이미지로 하여 진행하였다. 학습은 배치크기 2로 총 800 epoch을 수행하였으며 10 epoch 단위로 체크포인트를 저장하여 결과 이미지에 대한 비교 분석을 수행하였다.

Table 1. Training Parameters Settings

Parameter	Value
Batch Size	2
Learning Rate	0.0002
Epochs	800
Attention Heads	8
Hidden Dimension	512

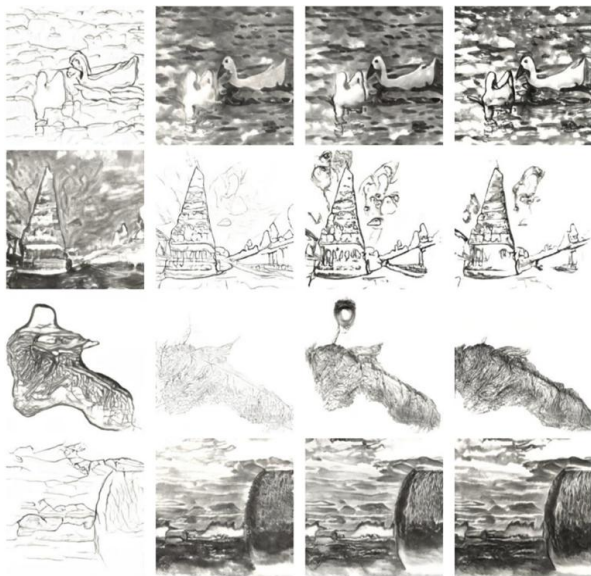


Fig. 2. Generated Images by Training Epochs

[그림 2]에서는 학습 진행에 따른 생성 이미지의 변화를 보여준다. 가장 왼쪽부터 50, 430, 700, 800 epoch의 훈련을 수행하였을 때의 생성 이미지이다. 50 epoch에서는 기본적인 형태만 나타나거나 형태가 무너지는 모습을 보인다. 430 epoch에서 수목화의 특징이 잘 드러나는 이미지를 생성할 수 있었다. 700 epoch 이후의 과도한 표현 문제는 과적합으로 인한 것으로 분석되며, 이는 early stopping 기법을 통해 해결할 수 있다. 특히 농담의 미세한 변화가 과장되어 나타나는 현상은 attention weight의 급격한 변화에서 기인하는 것으로 판단되며, learning rate scheduling과 gradient clipping을 통해 완화할 수 있다.



Fig. 3. Comparison of NST models

[그림 3]에서는 원본 이미지에 대한 AdaIN, CycleGAN, 그리고 제안된 모델의 스타일 변환 결과를 비교하였다. AdaIN은 ImageNet으로 사전 학습된 모델을 사용하였는데, 우수한 형태 보존력을 보여주었다. CycleGAN은 한국 전통 수목화 데이터셋의 데모 모델을 사용하였으며 데이터셋의 효용성과 CycleGAN 구조의 가능성을 확인할 수 있었다. 제안된 모델은 두 모델의 장점을 결합하여 완성도 높은 수목화를 생성할 수 있도록 설계되었다.



Fig. 4. Example of Traditional Korean Ink Painting

[그림 4]는 소정 변관식의 "관폭도"(1960년대 후반)로, 한국 전통 수묵화의 대표적 특징들을 잘 보여준다. 이 작품에서 나타나는 농담의 자연스러운 변화와 의미있는 여백의 활용은 본 연구에서 목표로 하는 수묵화의 핵심 요소들이다. 특히 폭포수의 역동적 표현에서 보이는 한 획 내의 미세한 강약 변화와 화면 좌측의 여백이 만들어 내는 공간감은 전통 수묵화만의 독특한 미학을 보여준다. [그림 3]의 비교 결과에서 볼 수 있듯이, 제안된 모델은 이러한 전통 수묵화의 특성을 성공적으로 재현한다. AdalN의 경우 형태는 보존하지만 수묵화 특유의 농담 변화가 부족하고, 기존 CycleGAN은 전체적인 수묵화 느낌은 구현하지만, 세밀한 붓터치의 강약이 손실된다. 반면 본 연구의 Self-Attention 기반 모델은 [그림 5]에서 보이는 것과 같은 자연스러운 농담의 흐름과 여백의 조화를 효과적으로 표현하여, 실제 전통 수묵화의 미학적 가치에 가장 근접한 결과를 생성한다.

### 3.6 Web Service

본 논문에서 제안한 시스템을 포함한 다양한 스타일 변환 모델로 한국화 변환을 누구나 쉽게 활용할 수 있도록 웹 서비스를 구현하였다.

프론트 엔드는 React와 NEXT.js 기반으로, 사용자 친화적이며 깔끔한 인터페이스를 제공한다. 웹 서버는 Flask 기반으로 구축하여 안정적으로 동작할 수 있도록 하였다. Flask는 PyTorch 모델을 효과적으로 서빙할 수 있는 가벼운 구조를 제공할 뿐 아니라, 서버 이전과 같은 서버 환경의 변화에도 유연하게 적용될 수 있는 특징을 가진다.

서버의 이미지 처리 과정은 세 단계로 나뉜다. 먼저 업로드된 파일의 유효성 검사를 수행한 뒤 크기 조정과 포맷 변환 등의 전처리를 수행한다. 이어서 지정된 모델에 입력해 스타일 변환한 뒤 결과 이미지를 후처리하여 클라이언트로 전송한다.

마지막으로, 다양한 모델을 선택하여 사용할 수 있도록 모델을 체계적으로 관리하고 그에 맞는 의존성을 유연하게 대처할 수 있도록 독립적인 환경을 구성하여 확장성을 높이고자 하였다.

해당 웹 서비스는 향후 사용자 피드백을 반영하여 지속적인 개선을 통해 서비스의 완성도를 높여나갈 예정이다.

## Generate Korean Painting

Hanji makes it simple for you to generate cool looking Korean Art in seconds, completely for free.

Generate your Art Learn more



Fig. 5. Main page of Web Generate Korean Painting

## IV. Conclusions

본 논문에서는 Self-Attention 매커니즘을 도입한 간소화된 CycleGAN 기반 스타일 변환 모델을 개발하였다. 기존 서양화 중심 스타일 변환 연구의 한계를 극복하고 한국 회화의 특성을 구현하였다는 점에서 의의가 있다.

또한 스타일 변환 기술의 실용적인 활용을 위해 웹 기반 서비스를 구현하였으며, 이를 통해 대중들이 손쉽게 한국화 스타일 변환에 관심을 가지고 경험할 수 있도록 하였다. 이는 향후 한국 회화 양식에 대한 연구에 활용되거나 전통 미술의 대중화에 기여하는 등의 다층적 의의를 가진다.

하지만 본 연구는 몇가지 한계를 가진다. 첫 번째로, 풍속화나 민화, 풍속화, 화조화 등 다양한 회화 양식에 적용하기 어려워 범용적 프레임워크를 개발하고 제한된 데이터로도 학습할 수 있는 기법을 연구할 예정이다. 두 번째로, 훈련 데이터셋에 포함된 형식의 그림과 상이한 구조의 이미지를 입력할 경우 생성된 이미지에서 원본 이미지의 구조가 유지되기 어렵다. 특히 난초화와 같이 여백 공간의 비중이 매우 높거나 단순한 획 구조를 가진 회화의 경우 생성 이미지의 품질이 저하되는 현상이 관찰되었다.

세 번째로, 기존의 Neural Style Transfer 평가 지표들은 주로 서양화 스타일 변환에 최적화되어 있다.

그러나 한국 수묵화와 같은 동양화는 여백의 활용과 추

상적 표현이 중요한 특징이므로, 기존 정량적 지표만으로는 문화적 특성과 예술적 가치를 완전히 평가하기 어렵다는 한계가 있다. 특히 수묵화의 핵심인 '농담의 자연스러움'과 '여백의 조화'는 수치적 측정보다는 시각적 평가와 문화적 맥락에서의 해석이 더욱 중요하다.

시각적 평가 결과, 제안된 모델이 기존 AdaIN과 CycleGAN에 비해 수묵화 고유의 특성인 붓터치의 농담 변화와 여백의 의미론적 활용을 더욱 효과적으로 표현하는 것을 확인하였다. 특히 Self-Attention 메커니즘을 통해 전역적 맥락을 고려한 스타일 변환이 가능해져, 수묵화의 철학적 가치까지 반영할 수 있었다.

향후 연구에서는 동양화 특성을 반영한 새로운 평가 지표 개발과 함께, 전문가 평가와 사용자 연구를 통한 종합적 성능 검증을 수행할 예정이다.

즉, 현재의 스타일 생성 모델의 구조로 수묵화의 특성을 완벽하게 재현하는 데 본질적인 제약이 있으며, 이를 극복하기 위한 새로운 기술적 패러다임이 필요함을 시사한다. 향후 연구에서는 다양한 양식의 전통 회화를 반영할 수 있는 기술의 개발로 확장될 수 있을 것으로 기대된다.

## REFERENCES

- [1] L. A. Gatys, A. S. Ecker and M. Bethge, "Image Style Transfer Using Convolutional Neural Networks," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 2414-2423, DOI: 10.1109/CVPR.2016.265.
- [2] H. N. L. Huy, L. V. Huy, and N. V. Hieu, "Artwork style transfer model using deep learning approach," *Cybernetics and Physics*, vol. 10, no. 3, pp. 127-137, 2021, DOI: 10.35470/2226-4116-2021-10-3-127-137.
- [3] Y. Jing, Y. Yang, Z. Feng, J. Ye, Y. Yu and M. Song, "Neural Style Transfer: A Review," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 11, pp. 3365-3385, 1 Nov. 2020, DOI: 10.1109/TVCG.2019.2921336.
- [4] S. Author, *Symbolism in Korean Ink Brush Painting*, e-book, 2006.
- [5] C. S. Kim, "The art class connected with the meaning of the blank space in the Korean painting," *J. Korean Child Art Educ.*, vol. 3, pp. 1-15, 2004.
- [6] J. An, S. Huang, Y. Song, D. Dou, W. Liu, and J. Luo, "ArtFlow: Unbiased Image Style Transfer via Reversible Neural Flows," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 862-871. arXiv:2103.16877.
- [7] Y. Deng, F. Tang, W. Dong, C. Ma, X. Pan, L. Wang, and C. Xu, "StyTr<sup>2</sup>: Image Style Transfer With Transformers," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 11326-11336.
- [8] J. E. Kyprianidis, J. Collomosse, T. Wang and T. Isenberg, "State of the "Art": A Taxonomy of Artistic Stylization Techniques for Images and Video," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 5, pp. 866-885, May 2013, DOI: 10.1109/TVCG.2012.160.
- [9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556, 2014.
- [10] C. Li and M. Wand, "Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 2479-2486, DOI: 10.1109/CVPR.2016.272.
- [11] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," arXiv:1603.08155, 2016.
- [12] X. Huang and S. Belongie, "Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 1510-1519, DOI: 10.1109/ICCV.2017.167.
- [13] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," arXiv:1607.08022, 2017.
- [14] V. Dumoulin, J. Shlens, and M. Kudlur, "A learned representation for artistic style," arXiv:1610.07629, 2016.
- [15] Y. Li, C. Fang, J. Yang, Z. Wang, X. Lu, and M.-H. Yang, "Universal style transfer via feature transforms," *Advances in Neural Information Processing Systems*, vol. 30, pp. 386-396, May 2017. arXiv:1705.08086.
- [16] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, vol. 27, pp. 2672-2680, 2014.
- [17] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2223-2232.
- [18] M. S. Ko, J. H. Kim, and B. Y. Kang, "GANMOOK: Generation of Korean ink-and-wash painting using generative adversarial networks," in *Proceedings of the Korea Information Processing Society Conference*, 2017, pp. 867-870.
- [19] C. Li and M. Wand, "Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 2479-2486, doi: 10.1109/CVPR.2016.272.
- [20] Z. Zou, T. Shi, S. Qiu, Y. Yuan and Z. Shi, "Stylized Neural Painting," 2021 IEEE/CVF Conference on Computer Vision and

- Pattern Recognition (CVPR), Nashville, TN, USA, 2021, pp. 15684-15693, doi: 10.1109/CVPR46437.2021.01543.
- [21] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 5998-6008.
- [22] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale," International Conference on Learning Representations (ICLR), 2021. DOI: 10.48550/arXiv.2010.11929.
- [23] S. Liu, T. Lin, D. He, F. Li, M. Wang, X. Li, Z. Sun, Q. Li, and E. Ding, "AdaAttN: Revisit Attention Mechanism in Arbitrary Neural Style Transfer," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 6649-6658. arXiv:2108.03647.
- [24] Y. Deng, F. Tang, W. Dong, C. Ma, X. Pan, L. Wang, and C. Xu, "StyTr<sup>2</sup>: Image Style Transfer With Transformers," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 11326-11336.
- [25] S. Yang, L. Jiang, Z. Liu, and C. C. Loy, "Pastiche Master: Exemplar-Based High-Resolution Portrait Style Transfer," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 7693-7702. arXiv:2203.13248.
- [26] M. Zhu, X. He, N. Wang, X. Wang, and X. Gao, "All-to-Key Attention for Arbitrary Style Transfer," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023, pp. 23015-23024.
- [27] Y. Deng, X. He, F. Tang, and W. Dong, "Z\*: Zero-shot Style Transfer via Attention Reweighting," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024, pp. 18502-18511.
- [28] X. Gao and Y. Zhang, "SRAGAN: Saliency Regularized and Attended Generative Adversarial Network for Chinese Ink-wash Painting Style Transfer," Pattern Recognition, vol. 162, Article 111344, 2025. arXiv:2404.15743.
- [29] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1125-1134. DOI: 10.48550/arXiv.1611.07004.

## Authors



Jin-Woo Cha received the B.S. degree in Computer Information from Inha Technical College, Korea, in 2020. He was a member of the KMS Lab at Inha University. His research interests include computer vision and machine learning.



Kyu-Cheol Cho received the B.S., M.S. and Ph.D. degrees in Computer Science and Information Engineering from Inha University, Korea, in 2005, 2007 and 2013, respectively.

Dr. Cho joined the faculty of the Department of Computer Science at Inha Technical College, Incheon, Korea, in 2016. He is currently an assistant professor in the Department of Computer Science, Inha Technical College. He is interested in cloud computing, green IT and web programming.