Digital Humanities:Practice and Theory Summative Project – April 2025
A Style Comparison of Miyazaki and Takahata
Using Colour Features and Transfer Learning

1.Introduction

1.1 Overview

Hayao Miyazaki and Isao Takahata, the two principal directors of the renowned Studio Ghibli, have garnered significant international acclaim for their artistic contributions in animation industry. Miyazaki's works is frequently characterized by vibrant, imaginative, and dreamlike visual aesthetics, whereas Takahata's works typically exhibit a more realistic and nuanced visual style. This study aims to combine data science and digital humanities methods to discuss the differences between the two directors in the animation from a quantitative perspective, and use the deep learning transfer learning model to automatically classify the screenshots. In doing so, this project verifies that artistic style can be quantitatively analyzed digitally, and discusses its limitations and feasibility, providing a new interdisciplinary perspective to the scholarly examination of cinematic works.

1.2 Project Objectives

1. Color Style Analysis

By extracting the mean RGB values, hue, saturation, and dominant color palettes from screenshots, this study quantifies and compares the visual differences in the cinematography between two directors.

2. Model Construction and Training

Based on transfer learning using pre-trained convolutional neural networks, develope a binary classifier to automatically distinguish screenshots.

3. Model Optimization and Evaluation

Use freezing convolutional layers, partially unfreezing layers for fine-tuning, and applying L2 regularization to enhance model accuracy. The model's performance was evaluated on the test set, with further analysis of misclassified scenes.

4. Digital Humanities Inquiry

By integrating quantitative findings with artistic interpretation, this study examines the feasibility and limitations of animated style classification, offering insights for future interdisciplinary research.

2.background

2.1 Director Profile & Artistic Context

Hayao Miyazaki: The selected works — *Castle in the Sky (1986), My Neighbor Totoro (1988), Princess Mononoke (1997), Spirited Away (2001), The Wind Rises (2013)* — exhibit a distinctive visual style characterized by bright tones, high saturation, and a playful, fantastical quality, collectively evoking a warm and immersive atmosphere.

Isao Takahata: *The analyzed films — Grave of the Fireflies (1988), Pom Poko (1994), Only Yesterday (1991), My Neighbors the Yamadas (1999), The Tale of the Princess Kaguya (2013)* — demonstrate a more subdued, naturalistic palette, often employing muted colors, minimalist watercolor techniques, or ink-wash aesthetics, reflecting a restrained and contemplative visual approach.

Despite their long-term collaboration at Studio Ghibli and shared foundational artistic principles, Miyazaki and Takahata maintained markedly divergent stylistic signatures, making their filmography an exceptional case study for visual style analysis and automated classification.

2.2 Related Works & Digital Humanities Approach

This study employs the K-Means clustering algorithm to extract the dominant color palette from screenshots of each film, with five primary color swatches selected per movie to represent its overall visual style. As an unsupervised learning algorithm, K-Means partitions image pixels into distinct color clusters by minimizing the squared Euclidean distance between samples and their nearest centroid (Lloyd, 1982), thereby identifying the most representative color combinations.

To further quantify the visual distinctions between the two directors, this study adopts a transfer learning approach using a Convolutional Neural Network (CNN), with VGG16 as the pre-trained model for director-style classification. CNNs are a class of deep learning architectures widely used in image recognition, capable of automatically extracting hierarchical texture and color features from images (LeCun, Bengio, & Hinton, 2015). The VGG16 architecture, renowned for its structural simplicity and stable performance in small-sample tasks, has been established as a benchmark model for numerous computer vision applications (Simonyan & Zisserman, 2014).

2.3 Motivation

This study focuses on the underexplored field of "animation director style quantification", aiming to systematically analyze screenshots from Studio Ghibli films and train deep learning models to determine whether machines can distinguish between directors' styles based on color palettes and texture features in images. This approach not only enhances digital humanities research by providing deeper insights into animation aesthetics but also opens up the potential applications of image classification techniques in identifying artistic styles within cultural works.

3. Data Acquisition & Pre-processing

3.1 Data Collection

I took 50 images from each of the 10 films, making a total of 500 screenshots to form the original data set. The number of screenshots per director accounted for 50% of the dataset, and to ensure diversity and coverage, I used a random method of screenshots, roughly every two to three minutes, until the entire film.

3.2 Pre-processing

First, all images were uniformly converted to RGB format using the Python Imaging Library (PIL). For each image, we calculated the mean values of the red, green, and blue channels, then automatically annotated them with corresponding film titles and director labels.

To capture the overall color tone of each film, we employed the K-Means clustering algorithm to group pixels from all screenshots, extracting five dominant colors that collectively represent the film's unique color palette. For computational efficiency, images were uniformly downsampled to a lower resolution prior to processing.

Furthermore, we transformed the images from RGB to HSV color space to extract the average hue and brightness values for each image. This additional step enabled us to quantify subtle differences in color preferences and lighting effects across films.

3.3 Data Augmentation and Dataset Partition

Due to the small amount of raw data, the training set employs data enhancement techniques：

- Random rotation
- Horizontal and vertical shifting
- Brightness adjustment
- Shearing transformation
- Random scaling
- Horizontal flipping

All pixel values were normalized to the range [0, 1] and resized to 224×224 pixels to meet VGG16's input requirements.

The dataset was divided into training set, verification set and test set at an 8:1:1 ratio. The verification set and test set are only normalized without data enhancement, so as to ensure the objectivity of training and evaluation at each stage. The model's generalization capability was ultimately assessed on the untouched test set.

4. Methodology

4.1 CNN Architecture & Transfer Learning

Considering the small number of images in the dataset, it is easy to overfit when using complex network structures such as ResNet50. Therefore, the VGG16 model with simple structure and mature transfer learning cases (Simonyan & Zisserman, 2014) is selected as the pre-training model in this project. Take full advantage of its generic features obtained on the large-scale ImageNet dataset (Deng et al., 2009). First, VGG16 is used to load the weights trained on ImageNet, and the include_top=False parameter is used to remove the top fully connected part. Then all convolutional layers are frozen as feature extractors to train the initial model. The architecture was enhanced with a GlobalAveragePooling2D layer for dimensionality reduction, followed by ReLU activation and Dropout layers(Srivastava et al., 2014) for regularization, culminating in a sigmoid output layer for binary classification. In the subsequent model training, the parameters were adjusted by gradually unfreezing the convolutional layer at the end of the network and adjusting the learning rate, and the L2 regularization was added to the fully connected layer of the network to rebuild the model and adjust the parameters. Finally, the higher accuracy and more stable model was selected as the final model. Its parameters are saved as/content/drive/MyDrive/Colab Notebooks/dh project/director_style_model_ft1_l2h5

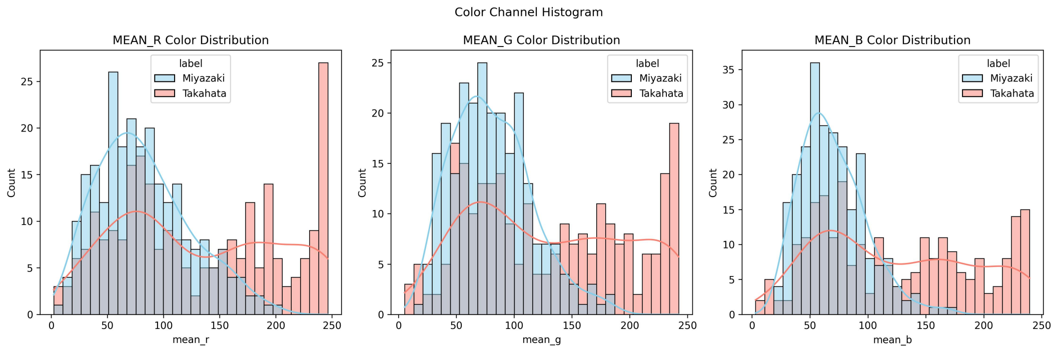
4.2 Experimental Design

During the initial training phase, the VGG16 convolution basis is frozen, only the new fully connected layer is trained, and the Adam optimizer (Kingma & Ba, 2015) is used to set the initial learning rate to 1e-4 to realize the adaptive update of parameters. The Early Stopping strategy (Prechelt, 1998) is introduced to stop the training when the performance of the validation set does not improve within several epochs to avoid unnecessarily long training and prevent the model from overfitting. The learning rate is reduced by ReduceLROnPlateau when the performance of the validation set stagnates, and the possibility of the model jumping out of the local optimal solution is improved. The model is trained for 15 epochs and saved according to the optimal performance of the validation set. Through the running results, it was found that the accuracy of the model still had room for improvement, so multiple rounds of parameter tuning were carried out until the model was seriously overfitting. In order to further suppress overfitting, L2 regularization is introduced in the final model construction, and Fine-tuning is performed at the back end of the network to gradually unfreeze the last several convolutional layers of VGG16.

5. Results： Color analysis and model performance

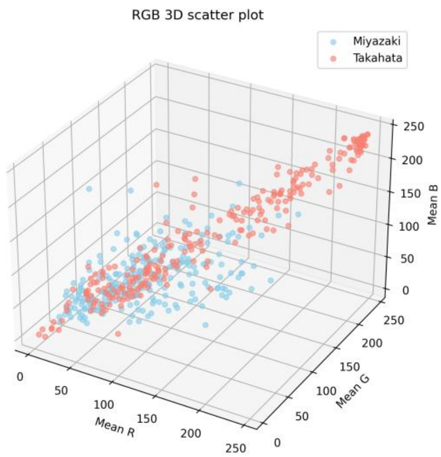
5.1 Colour Analysis

1. RGB & HSV histograms



The overall brightness distribution of Miyazaki's works is high, and the average RGB value is mostly concentrated in the middle and high range. Takahata's works are more widely distributed, and the frequency of low-brightness segments is significantly higher. In terms of hue/saturation, Hayao Miyazaki is more bright and high color contrast, while Isao Takahata's clips often show simple colors such as gray brown and light ink.

2. 3D RGB scatter plot



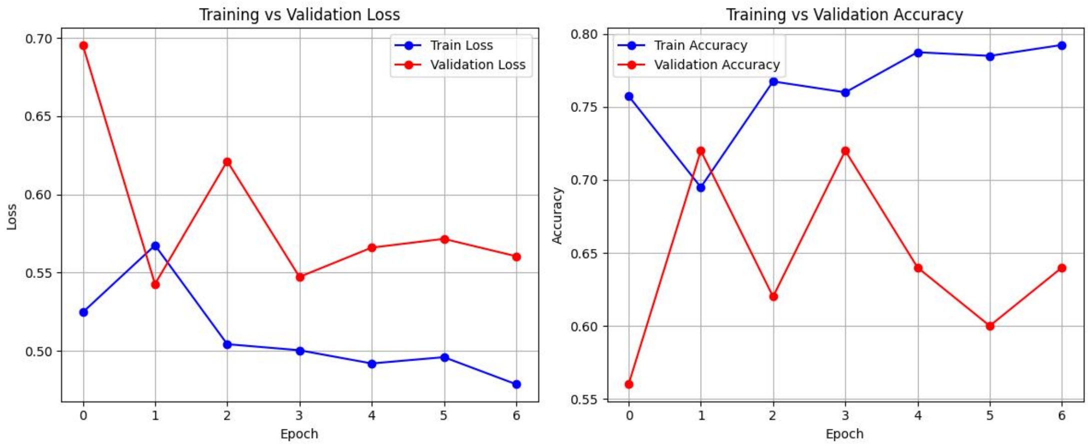
Miyazaki's work shows a wider range of variations in color brightness and tone, reflecting the rich, fantastical and emotional visual effects of his work. Takahata's work is more uniform and stable in overall color style, presenting a natural and realistic effect with attention to detail.

3. Color palette visualization

Through the analysis of color distribution and main color, it can be seen that the overall color of Miyazaki Hayao's works is brighter and more saturated, which reflects the construction of fairy tale and fantasy world. However, Isao Takahata's works are more natural and simple in tone, especially in "Only Yesterday" and "Grave of the Fireflies", which reflect the visual language of life and realism. After the 1990s, Miyazaki Hayao's works became more colorful, and Isao Takahata gradually integrated bright gray scale colors in his style exploration.

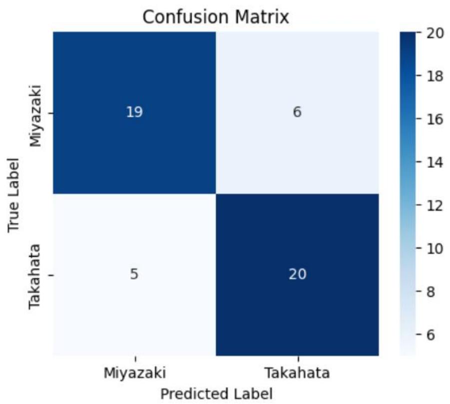
5.2 Model Performance

1. Train & validation curves



From the training loss curve, we can see that the training loss of the model continues to decrease as the number of training epochs increases, indicating that the model is constantly learning and optimizing its fit to the training data. The validation loss curve also shows a similar downward trend, indicating that the performance of the model on unseen data is also steadily improving, which r eflects the good generalization ability of the model. By comparing the loss and accuracy curves of training and validation, it can be observed that the trend of change is relatively similar. In the initial epochs, the training accuracy and validation accuracy steadily improve and the gap between them is small, which means that the model does not overfit significantly. However, as epochs increase, the performance on the validation set becomes unstable and slightly overfitting occurs. To prevent the model from falling into local optima or overfitting during training, I introduce Early Stopping and ReduceLROnPlateau. As can be seen from the graph, when the validation loss does not improve over multiple epochs, the system automatically reduces the learning rate and stops training when the validation performance reaches its best. This not only saves training time, but also ensures that the model parameters are saved for optimal performance.

Confusion Matrix & Accuracy



	precision	recall	f1-score	support
Miyazaki	0.79	0.76	0.78	25
Takahata	0.77	0.80	0.78	25
accuracy			0.78	50
macro avg	0.78	0.78	0.78	50
weighted avg	0.78	0.78	0.78	50

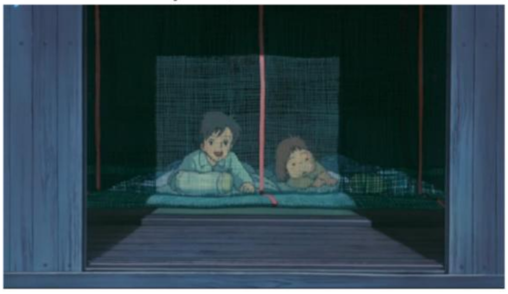
```
Confusion Matrix:
 [[19  6]
 [ 5 20]]
```

The overall accuracy of the model is about 78%, and the performance is balanced across all metrics, although there are still about 22% misclassification, but there is no obvious bias or major misclassification. From the overall results, the model can well learn and distinguish the visual styles of most of the works of the two directors. The precision and recall values are between 0.76 and 0.80, indicating that the model does not have serious missed detection or false detection for the prediction of the two categories, and this relatively balanced precision and recall rate also makes the F1-score at a medium-high level of about 0.78. In the confusion matrix, 19 Miyazaki classes were correctly predicted and 6 Miyazaki classes were misclassified. The Takahata category has 20 correctly predicted and 5 misclassified. Although there are still some misclassification cases in local areas, there is no extreme classification bias in general, and the model maintains a relatively balanced performance in distinguishing the works of the two directors.

5.3 Error Analysis

File: miyazaki/2-27.png, True: Miyazaki, Predicted: Takahata

True: Miyazaki, Predicted: Takahata



File: miyazaki/2-8.png, True: Miyazaki, Predicted: Takahata

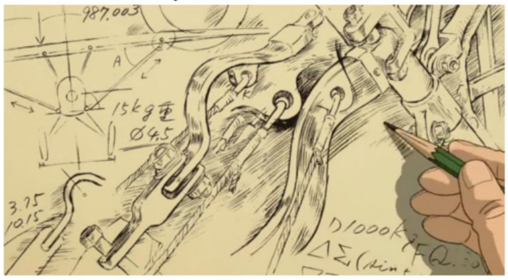
True: Miyazaki, Predicted: Takahata

File: miyazaki/2-9.png, True: Miyazaki, Predicted: Takahata

True: Miyazaki, Predicted: Takahata

File: miyazaki/5-14.png, True: Miyazaki, Predicted: Takahata

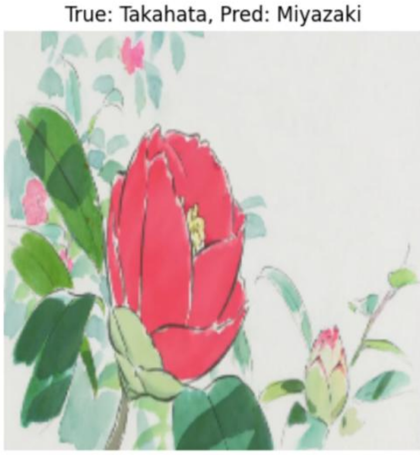
True: Miyazaki, Predicted: Takahata



Through case studies of misclassified samples, we observed that Miyazaki's and Takahata's visual styles occasionally exhibit overlapping characteristics, leading to model confusion when distinguishing between the two directors. Most misclassified frames displayed hybrid features—for instance, color brightness resembling Miyazaki's signature palette while compositional details aligned with Takahata's aesthetic, or atypical lighting/contrast that deviated from the film's dominant visual style. These ambiguities caused the model to assign incorrect labels due to conflicting feature representations.

True: Takahata, Pred: Takahata

True: Takahata, Pred: Miyazaki



True: Miyazaki, Pred: Miyazaki

True: Miyazaki, Pred: Takahata

In my analysis, we examined four representative misclassified frames (two per director). For Takahata's works, despite their distinct ink-wash aesthetics and deliberate white-space composition, one frame was incorrectly labeled as Miyazaki—likely due to unusually high color saturation, a trait more typical of Miyazaki's style. Conversely, two Miyazaki frames with fantastical hues and intricate details were misclassified as Takahata, potentially because subdued lighting or ambiguous scene layouts diluted their characteristic vibrancy.

6.Conclusion

6.1 Key Findings

This study employed digital humanities and data science methodologies to conduct a quantitative analysis of the distinctive visual styles of animation directors Hayao Miyazaki and Isao Takahata. By developing a VGG16-based transfer learning classifier, we systematically validated the statistically significant differences between their artistic approaches from both colorimetric and deep learning perspectives.

The experimental results demonstrated that under the condition of only training the last four layers and freezing most parameters of VGG16, the model achieves an overall accuracy of about 78% on the test set, and shows a relatively stable classification effect. Firstly, by freezing most of the parameters of VGG16, only a few four layers such as the newly added fully connected layer and Dropout layer are trained. With the help of the common features learned on large-scale data sets, the problem of overfitting under small samples is also effectively avoided. Secondly, L2 regularization, Early Stopping and learning rate decay strategies are adopted to stabilize the training process and ensure that the model has good generalization ability. Finally, the training process and model parameters are completely saved, and more powerful data support is provided by visualization tools such as training curve and confusion matrix. The model successfully identified distinguishable "color/texture" signatures between the directors from the limited data. Although there are still some misjudgment cases, the model still verifies that the animation director style has a certain degree of identifiability, and also shows the potential of deep learning in artistic style classification.

6.2 Limitations and Challenges

Although the research has achieved preliminary results, there are still the following shortcomings and limitations:

1. **Dataset Scale and Coverage**

   The current dataset suffers from limited sample size and insufficient diversity in frame selection, potentially compromising the model's generalizability when encountering edge cases or distinctive stylistic variations.

2. **Multidimensional Feature Representation**

   Animation style is not only reflected in the color statistics level, but also involves multi-dimensional information such as character modeling, storyboard composition, line texture and story plot. Under the condition of small sample size, the current CNN model is difficult to learn these deep semantic features systematically, and some error cases reflect this problem.

3. **Style Overlap and Consistency**

   Miyazaki Hayao and Isao Takahata often learn from each other and integrate their styles in Ghibli's works, there are large overlaps in the styles of some scenes, which increases the difficulty of model recognition and the risk of misjudgment.

7. Annex

	A	B	C	D	E	F
1	movie	label	hue	brightness	year	order
2	1Castle in the Sky	Miyazaki	0.42041912592421293	0.35913387450980394	1986	1
3	2My Neighbor Totoro	Miyazaki	0.3971488622992377	0.3432495607843137	1988	2
4	3Princess Mononoke	Miyazaki	0.437853211661724	0.2971520078431373	1997	3
5	4Spirited Away	Miyazaki	0.13787383139749249	0.4829886196078432	2001	4
6	5The Wind Rises	Miyazaki	0.2654182487384864	0.4504704705882353	2013	5
7	6Grave of the Fireflies	Takahata	0.2969910947279161	0.3162784705882353	1988	6
8	7Pom Poko	Takahata	0.37813999331065884	0.31327967843137255	1994	7
9	8Only Yesterday	Takahata	0.363488305137796	0.5254214431372549	1991	8
10	9My Neighbors the Yamadas	Takahata	0.23635171987343365	0.9157244862745099	1999	9
11	10The Tale of the Princess Kaguya	Takahata	0.30305498880572396	0.6778078901960786	2013	10
12						

34	/content/drive/MyDrive/Colab Notebooks/dh project	20.813759143675334	24.92299270757879	29.231975934634686	Miyazaki	1Castle in the Sky
35	/content/drive/MyDrive/Colab Notebooks/dh project	75.2983658362584	73.61376498376852	72.34433441942862	Miyazaki	1Castle in the Sky
36	/content/drive/MyDrive/Colab Notebooks/dh project	28.80048903356296	59.45673832819051	76.01029242475317	Miyazaki	1Castle in the Sky
37	/content/drive/MyDrive/Colab Notebooks/dh project	79.05805792204022	87.59488802186881	81.84710064944828	Miyazaki	1Castle in the Sky
38	/content/drive/MyDrive/Colab Notebooks/dh project	50.376143670657065	62.44174294236714	69.55240106223822	Miyazaki	1Castle in the Sky
39	/content/drive/MyDrive/Colab Notebooks/dh project	53.46404199045955	61.57246482897716	59.58010692317925	Miyazaki	1Castle in the Sky
40	/content/drive/MyDrive/Colab Notebooks/dh project	51.226728999697684	73.69908339509774	68.38351576500487	Miyazaki	1Castle in the Sky
41	/content/drive/MyDrive/Colab Notebooks/dh project	54.53711955935693	90.01119219068961	111.35127979245419	Miyazaki	1Castle in the Sky
42	/content/drive/MyDrive/Colab Notebooks/dh project	123.96321845834144	131.92601743596876	126.5172487271475	Miyazaki	1Castle in the Sky
43	/content/drive/MyDrive/Colab Notebooks/dh project	40.93849929274578	31.639050730891544	41.426325975725376	Miyazaki	1Castle in the Sky
44	/content/drive/MyDrive/Colab Notebooks/dh project	62.02318935464861	96.17975809262366	119.60478181741165	Miyazaki	1Castle in the Sky
45	/content/drive/MyDrive/Colab Notebooks/dh project	37.949730574099	66.36079959038048	87.4653257671634	Miyazaki	1Castle in the Sky
46	/content/drive/MyDrive/Colab Notebooks/dh project	58.609910368702934	59.68456210701027	53.59476640015628	Miyazaki	1Castle in the Sky
47	/content/drive/MyDrive/Colab Notebooks/dh project	109.79883682706841	73.48053250589628	59.213578833192855	Miyazaki	1Castle in the Sky
48	/content/drive/MyDrive/Colab Notebooks/dh project	132.46607853233007	83.87187527357817	51.50637409320428	Miyazaki	1Castle in the Sky
49	/content/drive/MyDrive/Colab Notebooks/dh project	82.59367019863815	108.72711959094221	122.16618232510658	Miyazaki	1Castle in the Sky
50	/content/drive/MyDrive/Colab Notebooks/dh project	75.1000163999109	87.07816175220962	87.82413680128849	Miyazaki	1Castle in the Sky
51	/content/drive/MyDrive/Colab Notebooks/dh project	26.649144755762876	62.145364024080756	85.04731320956066	Miyazaki	1Castle in the Sky
52	/content/drive/MyDrive/Colab Notebooks/dh project	62.22751441754198	77.92238460183384	78.9755005880898	Miyazaki	2My Neighbor Totoro
53	/content/drive/MyDrive/Colab Notebooks/dh project	54.76656281135122	49.8362260366922	40.719961048888585	Miyazaki	2My Neighbor Totoro
54	/content/drive/MyDrive/Colab Notebooks/dh project	33.912514884640046	71.14220201074889	91.80101055788731	Miyazaki	2My Neighbor Totoro
55	/content/drive/MyDrive/Colab Notebooks/dh project	41.92629075302469	91.53118418944132	115.45197861160175	Miyazaki	2My Neighbor Totoro
56	/content/drive/MyDrive/Colab Notebooks/dh project	78.47901115077585	127.03129060556995	129.3259634390424	Miyazaki	2My Neighbor Totoro
57	/content/drive/MyDrive/Colab Notebooks/dh project	59.447083824997165	91.89182898235269	82.13114837392958	Miyazaki	2My Neighbor Totoro
58	/content/drive/MyDrive/Colab Notebooks/dh project	89.12424718554767	106.57037639205295	97.33962325450047	Miyazaki	2My Neighbor Totoro
59	/content/drive/MyDrive/Colab Notebooks/dh project	61.4320668198132	67.14484945434874	50.589959589956386	Miyazaki	2My Neighbor Totoro
60	/content/drive/MyDrive/Colab Notebooks/dh project	111.50120199635609	98.98906789453909	91.18223519570103	Miyazaki	2My Neighbor Totoro
61	/content/drive/MyDrive/Colab Notebooks/dh project	83.22687990816279	78.07657310419911	64.87265178876449	Miyazaki	2My Neighbor Totoro
62	/content/drive/MyDrive/Colab Notebooks/dh project	88.69068839053743	81.12277530299446	79.26205391573902	Miyazaki	2My Neighbor Totoro
63	/content/drive/MyDrive/Colab Notebooks/dh project	21.78613779813489	30.602328043765965	30.781841015465343	Miyazaki	2My Neighbor Totoro

Note: the form a complete reference/content/drive/MyDrive/Colab Notebooks/dh project/image_color_features CSV

8. Bibliography

- Lloyd, S. (1982). Least squares quantization in PCM. IEEE Transactions on Information Theory, 28(2), 129–137.

- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.

- Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.

- Wikipedia (n.d.) ‘Hayao Miyazaki’.[online]Available at: https://en.wikipedia.org/wiki/Hayao_Miyazaki (Accessed: 03 March 2025)

- Wikipedia (n.d.) ‘Isao Takahata’.[online] Available at: https://en.wikipedia.org/wiki/Isao_Takahata (Accessed: 01 April 2025).

- Wikipedia (n.d.) ‘Ghibli’.[online] Available at: https://en.wikipedia.org/wiki/Ghibli (Accessed: 03 March 2025).

- Napier, S. J. (2018). Miyazakiworld: A Life in Art. Yale University Press.

- McCarthy, H. (1999). Hayao Miyazaki: Master of Japanese Animation. Stone Bridge Press.

- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A Large-Scale Hierarchical Image Database. 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 248–255. Available at: https://doi.org/10.1109/CVPR.2009.5206848

- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research, 15(1), 1929–1958.

- Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference on Learning Representations (ICLR).

- Prechelt, L. (1998). Early Stopping — But When? In G. Orr & K. Müller (Eds.), Neural Networks: Tricks of the Trade (pp. 55–69). Springer.