The Study of Render Farm Image Classification Using Deep Neural Network

Mujiono Sadikin and Aulia Permata Sari

***Abstract* — Rendering on the making of movie animation is a process of combining the imagination of animators and artistic creativity by turning the graphic display into millions of moving images. To perform the rendering process, the render farm is used, which is a combination of a group of high-performance computers (super-computers) commonly referred to as Computer Generated Imager. Throughout this time, the render artist spends a considerable plenty of time on doing image analysis from render farms manually. Reflecting on that, the study presented in this paper propose a deep learning method can separate images of the results of render farms more quickly and accurately. The technique proposed in this study is the classification of approved or revised images as the results of the render farm machine by using Deep Neural Network (DNN) technique. To verify the DNN that best fits the dataset, experiments were carried out on several layer depths and adjustment of epoch. In terms of treatment of the dataset, the experiment scenario selected was percentage and cross validation. The best performance in the experiment results is provided by layer depth configuration of 10, epoch value 100. The configuration provides the value of accuracy, precision and recall of 90%, 96% and 92% respectively.**

***Keywords* — Deep Neural Network; Deep Learning; Image Classification; Render Farm.**

I. Introduction

The increasing development of the multimedia world and the growing consumer interest in animation movies has triggered competitions among many animation studios in producing quality movies. This condition has an impact on the high level of production and the creation of new ideas amidst the creative animation industry [1]. Such development can be seen from the increasing quantity of animated films spread on social media such as Youtube, Instagram and Vimeo [2]. In Indonesia, several studios have come out producing animation movies and been able to compete with the quality of animation from abroad. Some of the animation movies include Nussa, ASJ, Kiko, Zakstorm, and various other movies. Unfortunately, the process of making animation movies in Indonesia is considered requiring a longer time in the making to get the maximum results.

The longer process of making animation movies in Indonesia is due to several factors, including the lack of human resources and the ineffectiveness in several stages of the process carried out. In the process of making animation movies, the studio must go through 3 stages. These stages are pre-production, production, and post-production. Pre-production comprises some important parts for the continuation of the production process, which are storyboarding, visual style, character design, environmental and location design, property design, voice casting, and story real or animatic aspect. The second stage is the production process. At this stage, all of the elements established in the previous stage are constructed, including the making of layout setting, animation movement, character finalization, final layout, effects, matte painting, composition, and lighting. The last step taken is post-production. After passing the production stage and the film is ready to undergo rendering, the animated film is polished to produce a final display ready to be distributed [3]. The rendering process is one of the important keys in determining the quality of a film. Details of texture, lighting and three-dimensional effect will determine the image quality during the rendering process [4].

As presented in [5], rendering is the final process of making 3D animation, combining the imagination of animators and artistic creativity by making stories into millions of images and combining which into a video. The rendering results will determine the final visual effect and the quality of a film. For almost the past three decades, the machines used for rendering have developed very rapidly and complex, including the renderMan owned by Autodesk 3Ds Max and the render farm. However, the current most-frequently used rendering machine is the render farm. Render farm is a high performance computer system, built using clustering computer technique, in which clustering computers means combining as many computers as possible to maximize performance in processing images and graphics to be produced [4]. The computers are designed to make CGI or commonly called Computer Generated Image, which is usually used to produce films and visual effects as if they were real.

However, the images produced by render farm are not all pursuant to user wishes. There are times when some defective or damaged images are found, for example, color differences or human errors during the production process. In consequence, manual work is done to separate images that match the user criteria. This manual work is very tedious and ineffective as it is done one by one for many images. To improve this obstacle, an image classification process from the render farm is proposed.

Various classification techniques applied to image dataset have been studied and developed. The first study was the classification of dogs and cats by Liu Bang *et al.* [6]. Liu Bang and the team studied the classification of dogs and cats based on the previous study conducted by Kaggle. The study began with the issue of solving CAPTCHA which is very easy for humans but difficult for machines to process it

Submitted on July 16, 2022.
Published on August 31, 2022.
M. Sadikin, Universitas Bhayangkara Jakarta Raya, Indonesia.
(corresponding e-mail: mujiono.sadikin@gmail.com;
mujiono.sadikin@dsn.ubharajaya.ac.id)
A. Permata Sari, Universitas Mercu Buana, Jakarta.
(e-mail: auliaalpapermatasari@gmail.com)

automatically. The dataset used was the Asirra dataset provided by Microsoft Research totaling 2500 training images and 1250 testing images. The method in this classification study applied two approaches. The first approach was human-crafted features, Dense-SIFT, and SVMs (Support Vector Machines) in the classification process carried out. However, the best accuracy obtained from the first approach was 71.47% which was considered unsatisfactory. To get better performance accuracy, researchers applied the second approach which was in the form of SVM and the CNN (Convolutional Neural Network) algorithm. The best accuracy obtained from the second approach was 94.00% meaning that it was accurate enough to be implemented on other data.

The second study was the classification of leaf image to find out 13 diseases in plants [7]. The dataset used was leaf image data obtained from the internet. The images were grouped into fifteen different classes. Thirteen classes represented the disease-infected state and the other two classes represented healthy plants. The image dataset was then converted to a size of 256×256 pixels using OpenCV python to make them similar. The highest accuracy obtained in this study by Srdjan Sladojevic, et al was 96.3% after the 100^{th} training. Convolutional neural network learning has proven to be one of the most powerful machine learning methods in image classification. The accuracy obtained exceeds other classification methods. The convolutional neural network process can simplify images containing millions of pixels into a small feature set, thereby reducing the dimensions of data input while maintaining their most important variable [8].

The next study was the segmentation of brain images by Annisa Wulandari, et al [9]. The background of this study was the need for a certain method that can do segmentation on MRI images of brain tumors appropriately and automatically. The method used was in the form of segmentation with thresholding. The initial stage carried out in the research was preprocessing MRI images using median filtering to improve image quality. Furthermore, MRI images were segmented using thresholding to identify and separate the desired object from the background based on the distribution of gray level or image texture. Then watershed segmentation and cropping were carried out to separate skull bones from brain tissue. The last stage was to calculate the percentage of brain tumor area. The output of this brain study was the percentage of brain tumor area that can be useful as an analysis material for radiologists. From the system experiment it was found that the calculation of tumor area had an error average of 10%, while from the calculation experiment of the brain area, the error average was 6%.

Based on several studies above, this study is conducted to investigate an automation of image separation from the results of render farms that are fitted or not to the user's criteria. Before going through the classification process, preprocessing needs to be done first using OpenCV python tools. This tool functions as a modifier of image pixels to make them smaller so that the classification process can take place faster. After getting smaller pixels, the image is turned into a grayscale image. The grayscale image is then converted into a matrix line according to the degree of grayness of the image. The next stage is the classification process using the Deep Neural Network algorithm by changing the epoch and layer depth. Lastly, the calculation of accuracy, precision and recall is carried out. The best accuracy result obtained is 91% with epoch 100 using a depth of 10 layers.

II. Material And Methods

A. Methods

Machine learning (ML) is a general term referring to various algorithms that make predictions based on a collection of big data. The predicted dataset consists of millions of unique data points[10]. Branch of machine learning itself is deep learning, which is a set of methods used in high-level architecture. The origin of deep learning concept comes from artificial neural networks of the human brain [11]. The feature space used in the deep learning method is the layer. Layer is used for data modeling, analysis and decision making on classification. Deep learning has shown the success of solving problems in various existing image classification studies [12].

Of some existing studies, this study will develop different methods and case studies. The method used was an experiment by testing several DNN configurations. The focus point was one the data used, which was image data from the result of render farms and image classification process using the Deep Neural Network algorithm was carried out. The research limitation lies in the image data used, that was in the form of medium angle of an animated film with movements based on the needs of the film, the image used was in the .png format and the rendering resolution was 960×540 pixels. From the image data, then the label definition stage was carried out for the next classification process. Labels used in this study were divided into 2, namely approve and revise. approve is an image that is pursuant to the user wishes, and revise is an image that must be repeated as it is not pursuant to the user wishes.

B. Research Stages

The stages of research in this study were divided into three. The first was to get the best model using the deep neural network algorithm. The second is to make an epoch comparison from the best models obtained using a 6-layer deep neural network. The last was to test layer changes to calculate the accuracy, precision and recall obtained. The detailed description of the experiment scenario is shown in Fig.1 below.

The first step performed was to collect data by re-rendering all scenes and shots needed from the animated film. The data used was a collection of images from render farm owned by one of the students of product design at Mercu Buana University in their previous semester Final Project. The number of data obtained was 900 images with 200 images for data testing and 700 images for data training. The distribution of data training and testing was randomly selected to prove that the machine can learn well. The size obtained from the original image was 960×540 pixels with a standard of 25 fps (frames per second) and .png image format. Some examples of images used are:

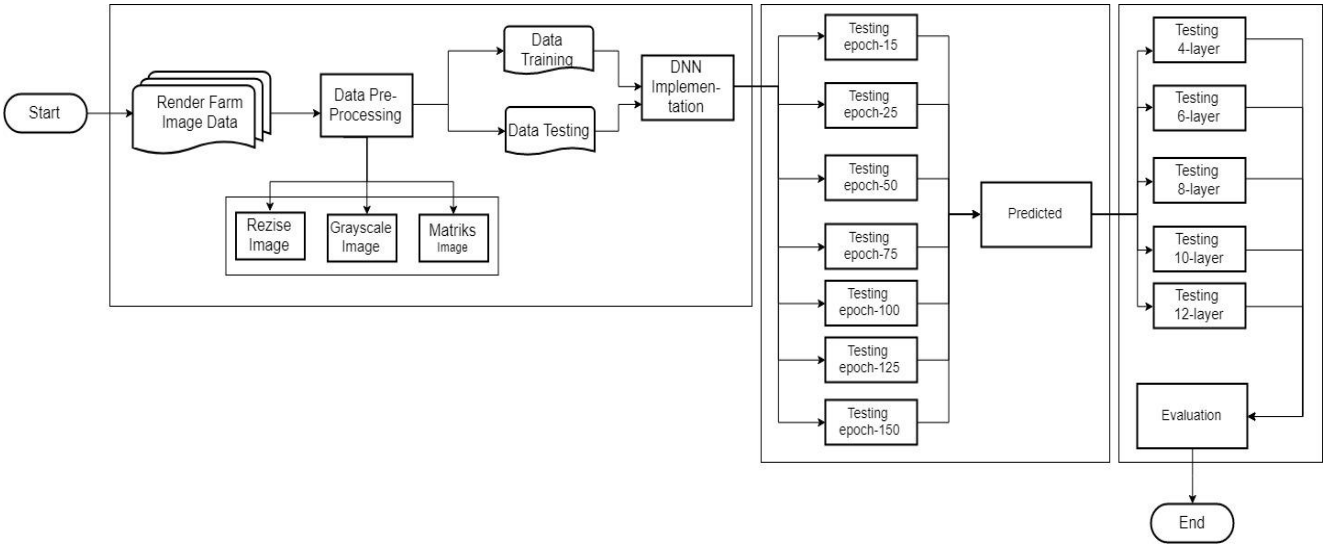

Fig. 1. Research Stages.

Fig. 2. The sample of image dataset.

C. Experiment Scenario

The experiment scenario conducted in this study are divided into three steps, which are as follows.

a. Adjustment/modification of epoch

Epoch training is a cycle of training for the entire dataset which has no standard rule for the number selected [13]. In this study, Epoch was changed 7 times, which were epoch 15, 25, 50, 75, 100, 125 and 150.

b. Adjust the hidden layer

Hidden layer, just like epoch training, does not have a definite answer for the application of the node used [14]. In this second experiment scenario, the experiment of change in the hidden layer was done five times. Changes to the selected hidden layer were 4-layer, 6-layer, 8-layer, 10-layer and 12-layer.

c. The treatment of the dataset

To validate the results obtained in the previous scenario, a cross validation experiment scenario was conducted. Cross validation is a widely used method to evaluate models, especially in data distribution [15]. In this third experiment scenario, there were 4 k-fold change experiments, namely 10-fold, 15-fold, 20-fold, and 25-fold that were applied to deep neural network (DNN) algorithms.

D. Pre-Processing Data

The process of preprocessing data is very important, because the raw data contains inconsistent values which will lead to producing data that do not meet the requirements [16]. The purpose of preprocessing is to improve the quality of the classification results. The steps in preprocessing the data in this study are:

a. Resize Image

The resize image step is about converting the image to a smaller pixel size using openCV python tool. Previous image data measuring 960×540 pixels was converted to a size of 50×50 pixels. These modified images were converted into pixels totaling 2500. The purpose of resize image is to speed up the process of converting pixel images into a row of matrices.

b. To convert image format to crayscale

After getting the image data with a smaller size, then the process of converting images to grayscale was done. The tool used was openCV python. OpenCV converts an image to grayscale image with the aim of making it easier to distinguish pixels using the degree of grayness of the image [17]. Table of degree of grayness of an image is shown in Fig. 3 below.

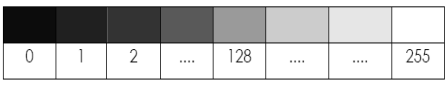

Fig. 3. Degree of Grayness of an Image.

c. Matrix Image

The final step of the preprocessing process was to turn the image pixels into a row of matrix numbers using openCV python. The aim of this process is to detect each pixel according to its visual content [18]. After the row of matrix is obtained, the next step is the implementation of the convolutional neural network algorithm.

E. Evaluation

The performance evaluation carried out in this study is based on the common criterias such us accuracy, precision and recall value of classification results. These performance criterias is computed using confusion matrix, which is a method that contains information about the actual class and the prediction of data classification results [16]. In this study, the approve label is a case considered positive, while the revise label is a case considered negative. The description of the definition of confusion matrix is presented as Table I.

TABLE I: CONFUSION MATRIX

Paramater	Remark
TP (True Positif)	Number of positive data considered true.
FP (False Positif)	Number of positive data considered false.
FN (False Negatif)	Number of negative data considered true.
TN (True Negatif)	Number of negative data considered false.

Each of performance criteria is computed using the formula below[13].

a. Accuracy

Accuracy is the number of classification ratios of positive data considered true and the total number of data:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}$$

b. Precision

Precision is the number of classification ratios of positive data considered true and the number of positive data considered true and false.

$$\text{Precision} = \frac{TP}{TP+FP}$$

c. Recall

Recall is the number of classification ratios of positive data considered true and the number of positive data considered true and negative data considered false

$$\text{Recall} = \frac{TP}{TP+FN}$$

III. RESULTS AND DISCUSSION

After preprocessing the data and implementing the convolutional neural network algorithm, this section will describe the results of experiments and analysis obtained. Several experiment scenarios have been carried out on the best models generated. The first experiment was to change the number of epoch training in the model with a 6-layer depth. From the results of the epoch training, the researcher made log loss percentage using the Tflearn python tool. Log loss is the value of probability of the comparison model, in which the smaller the prediction obtained, the higher the level of accuracy [19].

Table II presents log loss from the first experiment with 200 images of data testing. ID is the data testing image. From the results above, the log loss obtained was turned into the percentage form and divided into two categories. Category 1 is a percentage of less than 50%, and category 0 is a percentage of more than 50%. The next stage of the study was the calculation of accuracy, precision and recall using confusion matrix. Table III describes that the results of calculations obtained with the highest level of accuracy and recall is at epoch training 100 and 150. The best precision is obtained at epoch training 50.

The next experiment was to compare the changes of hidden layer from the results of the accuracy, precision and recall obtained in the previous stage. The results of accuracy, precision and recall obtained in the hidden layer stage are described in Table IV below.

From the experiment of changing the hidden layer, it is confirmed that the best accuracy and precision results were obtained at epoch training 100 and 10 hidden layers. The best recall was obtained from epoch training 150 and 8 hidden layers. These excellent results obtained shows that the more the number of hidden layers and epoch training used, the higher the level of accuracy obtained. The following are the results of the image and graphic classification with the best accuracy using DNN algorithm.

TABLE II: LOG LOSS EPOCH OF 6-LAYER DEPTH

ID	epoch-15	epoch-25	epoch-50	epoch-75	epoch-100	epoch-125	epoch-150
1	0.0001	0.0004	0.0001	0.0001	0.0014	0.0010	0.0027
2	0.8613	0.1455	0.0213	0.1370	0.0266	0.3559	0.8191
3	0.0003	0.0002	0.0003	0.0006	0.0015	0.0027	0.0032
4	0.8613	0.8710	0.8633	0.9098	0.8802	0.9526	0.9568
5	0.0057	0.0053	0.0008	0.0003	0.0020	0.0082	0.0088
6	0.0071	0.0122	0.0033	0.0018	0.0084	0.0132	0.0065
7	0.0133	0.0195	0.0087	0.0019	0.0078	0.3246	0.0073

TABLE III: RESULTS OF ACCURACY, PRECISION AND RECALL IN EPOCH COMPARISON

	epoch-15	epoch-25	epoch-50	epoch-75	epoch-100	epoch-125	epoch-150
Accuracy	88%	88%	89%	89%	90%	87%	90%
Precision	94%	96%	97%	96%	96%	93%	96%
Recall	91%	90%	90%	91%	92%	91%	92%

TABLEIV: RESULTS OF ACCURACY, PRECISION AND RECALL IN HIDDEN LAYER COMPARISON

Epoch	Hidden Layer	Accuracy	Precision	Recall
50	4-layer	90%	96%	92%
	6-layer	89%	97%	90%
	8-layer	90%	95%	92%
	10-layer	90%	97%	91%
	12-layer	90%	95%	93%
100	4-layer	90%	96%	92%
	6-layer	90%	96%	92%
	8-layer	89%	93%	93%
	10-layer	91%	98%	91%
	12-layer	88%	94%	91%
150	4-layer	90%	96%	92%
	6-layer	90%	96%	92%
	8-layer	88%	92%	93%
	10-layer	90%	96%	92%
	12-layer	90%	97%	91%

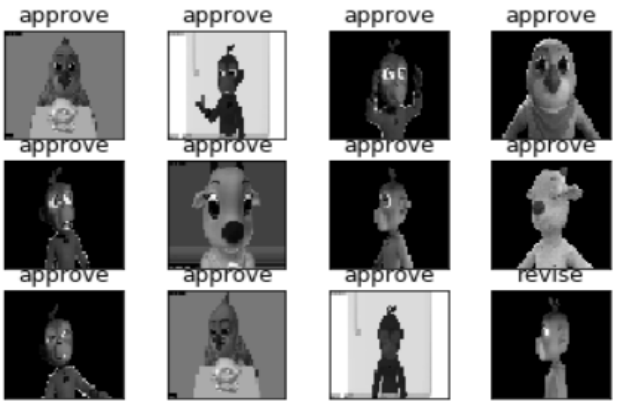

Fig. 4. Classification results with the best accuracy.

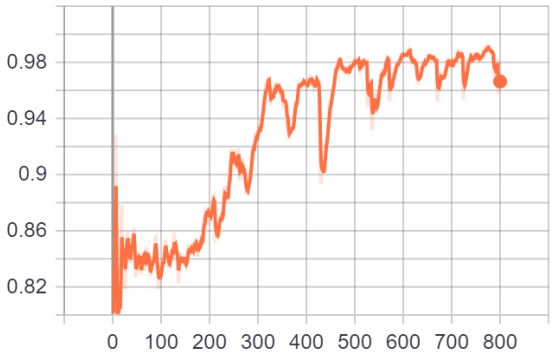

Fig. 5. Best accuracy graph with DNN algorithm.

The final stage was to change cross validation k-fold 4 times from the best accuracy results obtained in the previous experiment scenario. The results obtained are presented in Table V, which indicates that the best level of accuracy is obtained from 10-fold and 20-fold experiment with 90% of accuracy result.

TABLE V: RESULTS OF ACCURACY, PRECISION AND RECALL OF K-FOLD CHANGES IN EPOCH 100 AND 10 HIDDEN LAYERS

	10-fold	15-fold	20-fold	25-fold
Accuracy	90 %	89 %	90 %	89 %
Precision	96 %	95 %	96 %	96 %
Recall	92 %	91 %	91 %	91 %

IV. CONCLUSION AND FUTURE STUDY

In this study, the deep neural network algorithm implemented in the moving image data from the results of render farm obtained quite accurate results. The application of different number of epoch and layer depth influenced the results classification and level of accuracy obtained. The best result obtained during the experiment was the use of 10-layer depth with epoch 100. The dominant effect also lies in the pixel size of the image to be processed, where the greater its pixel image, the more the attributes, and in consequence the longer the classification process will take place. Suggestion for the development of future study is to do the classification on moving videos using a comparison of several different algorithms. This aims to simplify the process of making animation movie in the image classification that has been converted into a moving video.

REFERENCES

[1] Ma L, Qian C, Liu Z, et al. Exploring the Innovation System of the Animation Industry : Case Study of a Chinese Company. *Sustainability* 2018; 1–15.
[2] Muflihah A. Perancangan Sistem Rigging Pada Karakter Dalam Film Animasi "Iyatna".
[3] Rochman F, Subiyantoro H et. a. *Rencana Pengembangan Animasi Nasional 2015-2019*. 2015.
[4] Amann J, Wüthrich CA. Using Image Quality Assessment to Test Rendering Algorithms. In: *21st International Conference on Computer Graphics, Visualization and Computer Vision 2013*. 2013, pp. 205–214.
[5] Wu J, Xu H, Vassilev T. Design of 3D Rendering Platform Based on Cloud Computing. *2016 4th Int Conf Enterp Syst* 2016; 153–159.
[6] Liu B, Liu Y, Zhou K. Image Classification for Dogs and Cats. 1–10.
[7] Sladojevic S, Arsenovic M, Anderla A, et al. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. 2016.
[8] Zhang X, Zhang Y, Han EY, et al. Whole mammogram image classification with convolutional neural networks. In: *Proceedings - 2017 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2017*. 2017, pp. 700–704.
[9] Wulandari A, Sigit R, Bachtiar MM. Brain Tumor Segmentation to Calculate Percentage Tumor Using MRI.
[10] Nichols JA, Herbert Chan HW, Baker MAB. Machine learning: applications of artificial intelligence to imaging and diagnosis. *Biophys Rev* 2019; 11: 111–118.
[11] Ejbali R, Zaied M. A dyadic multi-resolution deep convolutional neural wavelet network for image classification. *Multimed Tools Appl* 2018; 77: 6149–6163.
[12] Rivenson Y, Zhang Y, Günayd H, et al. Phase recovery and holographic image reconstruction using deep learning in neural networks. Epub ahead of print 2018. DOI: 10.1038/lsa.2017.141.
[13] Ikonomakis EK, Kotsiantis S, Tampakas V. Text Classification Using Machine Learning Techniques. *WSEAS Trans Comput* 2005; 4: 966–974.
[14] Wang Y, Xu L. Research on text categorization of KNN based on K-means for class imbalanced problem. In: *Proceedings - 2016 6th International Conference on Instrumentation and Measurement, Computer, Communication and Control, IMCCC 2016*. 2016, pp. 579–583.
[15] Rabinowicz A, Rosset S. Cross-Validation for Correlated Data. *J Am Stat Assoc* 2020; 0: 1–14.
[16] Sadikin M, Afiandi F, Alfiandi F. Comparative Study of Classification Method on Customer Candidate Data to Predict its Potential Risk. *Int J Electr Comput Eng*; 8. Epub ahead of print 2018. DOI: 10.11591/ijece.v8i6.pp4763-4771.
[17] Sachdeva A, Kapoor R, Sharma A, et al. Categorical Classification and Deletion of Spam Images on Smartphones Using Image Processing and Machine Learning. *Proc - 2017 Int Conf Mach Learn Data Sci MLDS 2017* 2018; 2018-Janua: 23–30.
[18] Loussaief S, Abdelkrim A. Machine learning framework for image classification. *2017 Int Conf Inf Digit Technol* 2017; 58–61.
[19] Jiang Z, Gao S, Li M. An improved advertising CTR prediction approach based on the fuzzy deep neural network. *PLoS One* 2018; 13: 1–25.

Mujiono Sadikin is faculty member of Faculty of Computer Science Universitas Bhayangkara Jakarta Raya. He held doctoral degree from Universitas Indonesia, Jakarta 2017. His research area is in Data Mining, Machine Learning, and IT Governance as well. Some of his experiences are: As team leader in IT Governance and Procedure preparation of Directorate Land & Transportations Ministry of Transportation, Team leader of IT Audit and Assessment Universitas Mercu Buana, and some more.

Aulia Permata Sari is a final year undergraduate student in Informatics Engineering, Faculty of Computer Science, Universitas Mercu Buana Jakarta in 2019. She was born in Surabaya on December 15th, 1995. She has work experience as a technical artist for 5 years in an animation studio. She was interested in animation, multimedia, machine learning, and programming.