ISSN: 2150-8402 Volume 14, Number 4, November 2022

Scientific Research Publishing

Journal of Intelligent Learning Systems and Applications

ISSN: 2150-8402

https://www.scirp.org/journal/jilsa

Journal Editorial Board

ISSN Print: 2150-8402 ISSN Online: 2150-8410
https://www.scirp.org/journal/jilsa

Journal of Intelligent Learning Systems and Applications, 2022, 14, 43-131
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

Table of Contents

Volume 14 Number 4 November 2022

Journal of Intelligent Learning Systems and Applications (JILSA)
Journal Information

SUBSCRIPTIONS

The *Journal of Intelligent Learning Systems and Applications* (Online at Scientific Research Publishing, https://www.scirp.org/) is published quarterly by Scientific Research Publishing, Inc., USA.

Subscription rates:
Print: $89 per copy.
To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

SERVICES

Advertisements
Advertisement Sales Department, E-mail: service@scirp.org

Reprints (minimum quantity 100 copies)
Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.
E-mail: sub@scirp.org

COPYRIGHT

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: jilsa@scirp.org

Journal of Intelligent Learning Systems and Applications, 2022, 14, 43-56
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

Multi-Label Chest X-Ray Classification via Deep Learning

Aravind Sasidharan Pillai

University of Illinois Urbana-Champaign, Champaign, IL, USA
Email: pillai.aravinds@gmail.com

How to cite this paper: Pillai, A.S. (2022) Multi-Label Chest X-Ray Classification via Deep Learning. *Journal of Intelligent Learning Systems and Applications*, **14**, 43-56.
https://doi.org/10.4236/jilsa.2022.144004

Received: August 9, 2022
Accepted: October 29, 2022
Published: November 1, 2022

Abstract

In this era of pandemic, the future of healthcare industry has never been more exciting. Artificial intelligence and machine learning (AI & ML) present opportunities to develop solutions that cater for very specific needs within the industry. Deep learning in healthcare had become incredibly powerful for supporting clinics and in transforming patient care in general. Deep learning is increasingly being applied for the detection of clinically important features in the images beyond what can be perceived by the naked human eye. Chest X-ray images are one of the most common clinical method for diagnosing a number of diseases such as pneumonia, lung cancer and many other abnormalities like lesions and fractures. Proper diagnosis of a disease from X-ray images is often challenging task for even expert radiologists and there is a growing need for computerized support systems due to the large amount of information encoded in X-Ray images. The goal of this paper is to develop a lightweight solution to detect 14 different chest conditions from an X ray image. Given an X-ray image as input, our classifier outputs a label vector indicating which of 14 disease classes does the image fall into. Along with the image features, we are also going to use non-image features available in the data such as X-ray view type, age, gender etc. The original study conducted Stanford ML Group is our base line. Original study focuses on predicting 5 diseases. Our aim is to improve upon previous work, expand prediction to 14 diseases and provide insight for future chest radiography research.

Keywords

Data Science, Deep Learning, X-Ray, Machine Learning, Artificial Intelligence, Health Care, CNN, Neural Network

1. Introduction

Radiology is a branch of medicine that can be divided into diagnostic radiology and interventional radiology [1]. Diagnostic radiology involves examining the medical images to diagnose diseases and abnormalities. Chest X-ray radiography is the most common imaging examination that demands correct and immediate interpretation to avoid life-threatening diseases. The challenge arises when these images have to be interpreted by radiologists who are limited by speed, experience and the cost involved to get a certified radiologist. Therefore, health care industry turned towards deep learning algorithms to automate and generate accurate radiology reporting.

Deep learning in health care is bursting with possibility and remarkable innovation by providing the ability to analyze vast quantities of data at exceptional speed without compromising on accuracy [2]. It enables the creation of algorithms that can learn and make predictions. In contrast to rules-based algorithms, machine learning takes advantage of increased exposure to large data sets and possesses the ability to improve performance with such exposures and learn with experience.

Deep learning is a subset of artificial neural networks that are statistical and mathematical methods inspired by the way biological nervous system processes information with a large number of highly connected neurons, nodes or cells. Neural networks are structured as one input layer, one or more hidden layers and an output layer. Every hidden layer consists of a set of neurons that are fully connected to all neurons in the previous layer. The strength of such a connection is determined by weights of variable or features that associate inputs with outputs. For a neural network model to perform efficiently and accurately, these weights must be set to suitable values, which are estimated through training.

From deep learning perspective, radiology images need to be pre-processed differently due to variations in processors and memory restrictions. X-ray images in general are 2D images of a 3D human body. DL algorithms especially convolutional neural networks (CNN) have proved more successful in training the models with 2D images than 3D images that adds an additional dimensionality to the problem. Convolution is a mathematical operation that employs a type of filtering to determine the most useful features from a dataset, thereby having applications in finding patterns in signals or filtering signals.

1.1. Related Works

In recent years, large sets of radiology images have been made public. The availability of such datasets helped crowd-source the development and evaluation of deep learning models [3] [4] [5] [6]. In 2017, NIH Clinical Center released over 100,000 chest x-ray images, which comprises 108,948 frontal-view X-ray images of 32,717 unique to the scientific community. There were several studies based on this data. Wang *et al.* [7] through their paper "ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on weakly-Supervised Classification and Localization of Common Thorax Diseases" demonstrated that the commonly occurring thoracic diseases can be detected and even spatially located via a unified weakly supervised multi-label image classification and disease localization framework.

In 2019, MIT published MIMIC-Chest X-Ray Database (MIMIC-CXR) collection of more than 350,000 chest x-ray images associated with 227,943 studies. These images consist of both frontal and lateral views and were compiled from Beth Israel Deaconess Medical Center in Boston during years 2011 to 2016. Images are provided with 14 labels derived from free-text radiology reports using NLP tools [8].

CheXpert is another such dataset consisting of 224,316 chest radiographs of 65,240 patients who underwent a radiographic examination from Stanford University Medical Center between October 2002 and July 2017, in both inpatient and outpatient centers. Based on associated radiology reports, these X-rays images were labeled as positive, negative, or uncertain for the presence of 14 common chest radiographic observations. CheXpert data has attracted strong attention in building pre-trained learning models to address the challenges in X-Ray image processing, classification and segmentation. CheXNet is one prominent project by Stanford ML Group which is considered to have state of the art performance in classifying diseases even better than expert radiologists.

1.2. Baseline

In this paper, we take inspiration from state-of-the-art CheXNet model to train similar models based on CNN that perform multi-class, multi-label classification with transfer learning from models trained on Imagenet data. Irvin *et al.* [1], in their paper developed models to predict five lungs' conditions such as Atelectasis, Cardiomegaly Consolidation, Edema and Pleural Effusion. Their study achieved AUC ranging from 0.85 to 0.93. Our goal is to expand the predictions from 5 diseases to 14 diseases and achieve comparable AUCs, and also attempt to compare various pre-trained CNN models and their performance.

2. Approach and Methods

Our approach primarily focuses on developing CNN classifier to predict the labels. During research, it is a good practice to evaluate models on different datasets since deep learning algorithms are often validated on historical data, yielding guaranteed performance but are unable to attain same levels of accuracy when operating outside the training data range. This is called as overfitting. We employ a method to compare the performance of different models based on success metrics of the model on test data set of images and determine the best performing model under different parameter setting.

2.1. Convolutional Neural Network

CNN is a deep Learning algorithm which can take an input image, assign importance as weights to various features in the image and be able to differentiate one from the other. The main advantage of Convolutional Neural network is that it has the capability to capture the temporal and spatial dependencies in an image by applying relevant filters [9]. CNN architecture is composed of a convolutional layer, a pooling layer and a fully connected layer. The function of a convolutional layer is to extract features from images. Each convolutional layer can have multiple convolution kernels and the convolutional layer is calculated as follows:

$$x_j^{\,l} = f\left(\sum_{i \in M_j} x_j^{\,l-1} * k_{ij}^l + b_j^{\,l}\right),$$

where $x_j^{\,l-1}$ is the characteristic map of the output of previous layer, $x_j^{\,l}$ is the output of the *i*th channel of the *j*th convolutional layer and *f*(.) is called the activation function. M_j is the subset of input feature maps, k_{ij}^l is a convolutional kernel and $b_j^{\,l}$ is its corresponding weight. A pooling layer sits between two convolutional layers to reduce the feature map dimensions and still maintain the scale of the features to an extent. Mean pooling and Max pooling are two main pooling methods used. A fully connected layer integrates multiple image maps after they are passed through multiple convolutional and pooling layers to obtain high-level semantics used to determine the class labels.

All the experiments in our work involve millions of parameters for multiple features and multiple layers of neural networks. This requires us to use a faster processing machine with graphic processing units (GPU) and cloud architecture that have pre-configured drivers and come with popular Python packages to build neural networks.

2.2. Experimental Setup

We used AWS Deep Learning AMI (Ubuntu 18.04) for training and prediction. These AMIs are pre-configured with PyTorch 1.7.1 and Python3.7 (CUDA 11.1 and Intel MKL) along with other basic data science software like pandas, numpy, scipy etc. We used g4dn.2xlarge machine with 1 Tesla T4 GPU.

2.3. Dataset

We used a version of Chexpert data with down sampled resolution to train the model. The original data is 439 GB in size whereas the down sampled version is just 11 GB. We chose the down sampled version to save training time and computing power. Sample images are shown in **Figure 1**.

Each individual X-ray radiograph is represented by the image and a feature vector containing patient id, sex, age, study number, X-ray view-type, and a vector of fourteen expert-labeled observations. We clean the data to omit the sex and age features from the equation because we evaluate solely on the image.

Train Data Distribution

This subset of dataset is split into train and validation sets for model evaluation with a split proportion of 80:20. **Figure 2** and **Figure 3** denote the distribution of train data and train labels respectively. An unseen test data set is used to evaluate predictions of classes by the trained models.

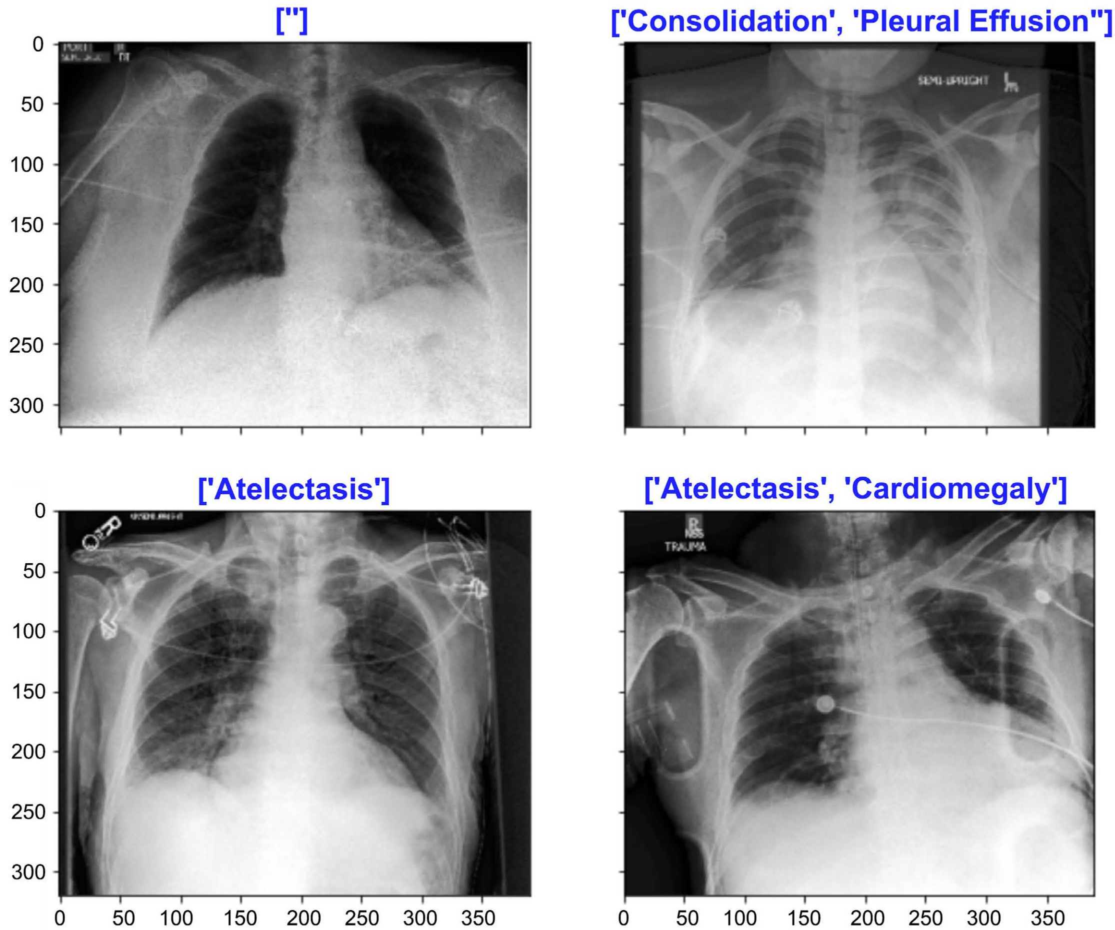


Figure 1. Sample images.

	Label	minusOneVal	oneVal	zeroVal	nanVal
0	No Finding	0	22381	0	201033
1	Enlarged Cardiomediastinum	12403	10798	21638	178575
2	Cardiomegaly	8087	27000	11116	177211
3	Lung Opacity	5598	105581	6599	105636
4	Lung Lesion	1488	9186	1270	211470
5	Edema	12984	52246	20726	137458
6	Consolidation	27742	14783	28097	152792
7	Pneumonia	18770	6039	2799	195806
8	Atelectasis	33739	33376	1328	154971
9	Pneumothorax	3145	19448	56341	144480
10	Pleural Effusion	11628	86187	35396	90203
11	Pleural Other	2653	3523	316	216922
12	Fracture	642	9040	2512	211220
13	Support Devices	1079	116001	6137	100197

Figure 2. Distribution of train data.

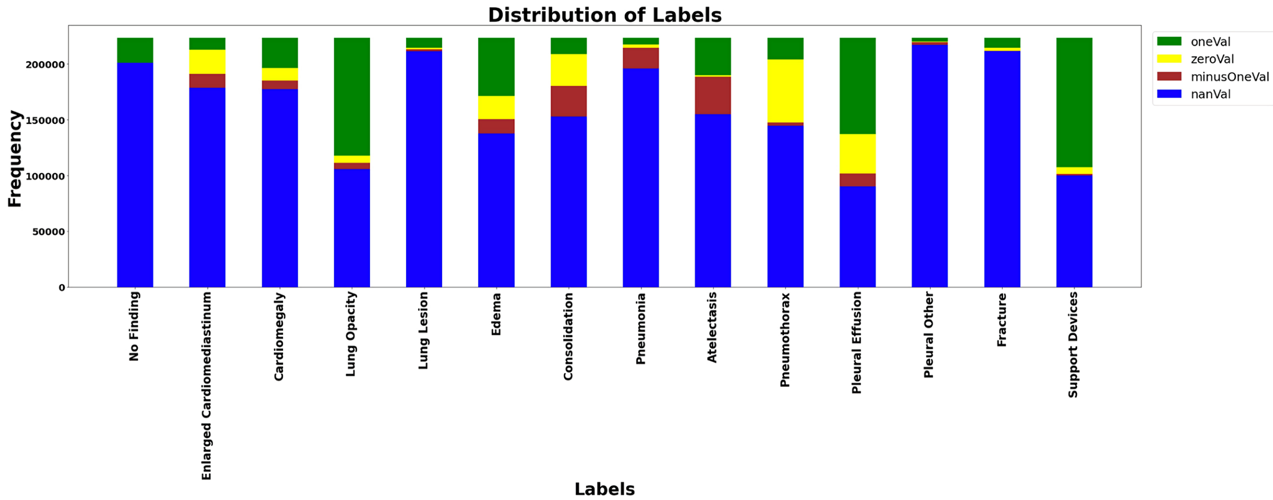


Figure 3. Distribution of train labels.

2.4. Data Uncertainty

Chest X-ray dataset contains images that are hard to classify for a certain disease to be present. So, radiologists often leave an uncertain label (value = −1) on such images. We followed approach similar to Irvin *et al.* [4] to handle uncertainties. Labels were split into u-zero and u-one categories based on previous studies. Category u-one means uncertain labels will be treated as positive and u-zero means negative. We considered labels Atelectasis, Edema as u-ones and rest of the labels as u-zero.

2.5. Data Pre-Processing

Dataset used for training and test are pre-processed through augmentation methods. This was done to increase the size and quality of the dataset. This process helps in solving problems related to overfitting and enhances the model's generalization ability during training and prediction of class labels.

2.6. Modeling

Based on our research, CNN architecture performs better on multi-class, multi-label classification of image dataset due to the reduction in number of parameters involved, without losing features that are critical for getting a good prediction. So, we investigated multiple models based on CNN architecture that will be discussed in detail further. Each of this model is run on train and test data with a batch size of 96 for up to 40 epochs, binary cross entropy as loss function, Adam optimizer with an initial learning rate of 0.001 which is multiplied by 10 each time the validation loss plateau after an epoch.

2.6.1. Custom Net (Simple Baseline Model)

We started with a simple custom-made CNN models. Trained the model from scratch where input is passed through 4 convolutional layers by random initialization and fine tuning the weights of all the layers. Each convolutional layer has a max pool layer and ReLU activations. Pooling layers used to reduce spatial volume. Figure 4 demonstrates list of parameters for each layer. Output of these layers is then passed to a fully connected sigmoid function to classify a collection of chest X-ray images. Sigmoid function helps to convert raw output values from classifier to corresponding probability values. We considered probability greater than 0.50 as positive detection.

2.6.2. DenseNet121

The core idea of DenseNet is to ensure maximum information flow between layers in the network by connecting all layers directly with each other. It has a stack of dense blocks followed by transition layers. Dense blocks contain different units such as convolutions, batch normalization and ReLU activations. Each dense block generates a fixed number of feature vectors which is called the Growth rate *i.e.*, the amount of information that layers can transmit. We trained DenseNet121 with initial weights from a pre-trained network on ImageNet data [10].

2.6.3. ResNet-50

ResNet-50 is a convolutional neural network that is 50 layers deep. We initialize the network with pre-trained weights, where knowledge is transferred from ImageNet data. Pre-trained network initial weights are frozen for first 6 layers used for feature extraction and only the weights of the last layer are adapted to re-train one or more layers with samples from the X ray dataset [11].

2.6.4. Inception_V3

This network has 48 layers depth that can make several improvements including label smoothing, factorized convolutions and uses an auxiliary classifier to propagate label information lower down the network. We initialize the network with

CustomNet

Modules	Parameters
ConvLayer1.0.weight	216
ConvLayer1.0.bias	8
ConvLayer1.1.weight	1152
ConvLayer1.1.bias	16
ConvLayer2.0.weight	12800
ConvLayer2.0.bias	32
ConvLayer2.1.weight	9216
ConvLayer2.1.bias	32
ConvLayer3.0.weight	18432
ConvLayer3.0.bias	64
ConvLayer3.1.weight	102400
ConvLayer3.1.bias	64
ConvLayer4.0.weight	204800
ConvLayer4.0.bias	128
ConvLayer4.1.weight	147456
ConvLayer4.1.bias	128
Lin1.0.weight	7168
Lin1.0.bias	14

Total Trainable Params: 504126

Figure 4. CustomNet training parameters.

pre-trained weights, where knowledge is transferred from ImageNet data and freeze these weights for first 8 layers [12].

2.6.5. Vgg16

The VGG network is a neural network that has already been pretrained on over a million images from the ImageNet database. The network has 41 layers. There are 16 layers with learnable weights, 13 convolutional layers, and 3 fully connected layers. One of the major disadvantages of the VGG16 Neural Network is the huge number of trainable parameters. It has more than 134 million trainable parameters. We did freeze the first 6 layers so as to limit the trainable parameters to 57 k [11] [13].

Figure 5 demonstrates trainable parameters for each of these models.

2.7. Success Metrics

Accuracy: It is the ratio of number of correct predictions to the total number of input samples.

Area Under Curve: AUC of a classifier is equal to the probability that the classifier will rank a randomly chosen positive example higher than a randomly chosen negative example [14].

F1 Score: F1 Score is the Harmonic Mean between precision and recall.

Precision: The precision is the ratio the number of true positives to the sum of true positives and false positives. The precision denotes the ability of the classifier not to label as positive a sample that is negative.

Recall: The recall is the ratio the number of true positives to the sum of true positives and false negatives. The recall denotes the ability of the classifier to find all the positive samples.

2.8. Model Training

We did a random 20% split of train data as validation data and trained model for a maximum of 40 epochs. ROC was used to determine early stopping criteria. Most of the models achieved best performance with 20 to 25 epochs. Figure 6 demonstrates variations in training metrics per epoch.

DenseNet model achieved highest training AUROC 78 and highest training accuracy 87% as indicated in Figure 7.

Training Parameters

	Model Name	Total Parameters	Trainable Parameters
0	CustomNet	504126	504126
1	DenseNet121	6968206	6968206
2	ResNet50	23772110	264078
3	Inception	21814254	28686
4	Vgg16	134317902	57358

Figure 5. Training parameters summary.

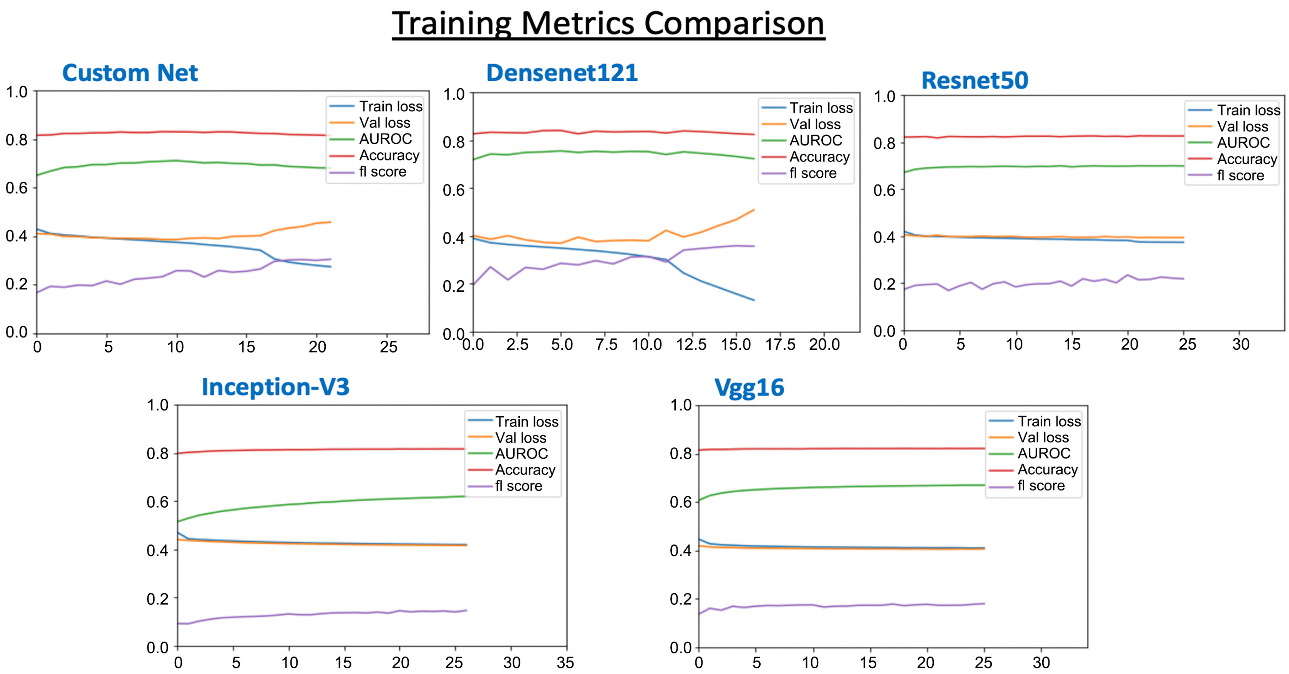


Figure 6. Training metrics per epoch.

2.9. Training and Validation Metrics

	Model Name	AUROC	Accuracy	f1 Score
0	CustomNet	0.740989	0.862637	0.259044
1	DenseNet121	0.779881	0.866486	0.271429
2	ResNet50	0.716902	0.854088	0.219912
3	Inception	0.650021	0.846562	0.151924
4	Vgg16	0.671470	0.848543	0.164998

Figure 7. Training metrics summary.

3. Results

To evaluate the performance of models, we used unseen test data to predict the multi classifications labels. The validation set contains 200 studies from 200 patients randomly sampled from the full dataset with no patient overlap with the train set. Test data consisted of 234 images. Distribution of test data and labels is shown in Figure 8.

3.1. Test Metrics

Model performance validated against success metrics. Details available are in Figure 9.

Based on overall test metrics, DenseNet121 achieved the best performance. This model achieved ROC score of 0.78 and accuracy of 87%. Other models ROC values ranged from 0.69 to 0.75 and accuracy from 83% to 86%.

Test Dataset

	Label	minusOneVal	oneVal	zeroVal	nanVal
0	No Finding	0	38	196	0
1	Enlarged Cardiomediastinum	0	109	125	0
2	Cardiomegaly	0	68	166	0
3	Lung Opacity	0	126	108	0
4	Lung Lesion	0	1	233	0
5	Edema	0	45	189	0
6	Consolidation	0	33	201	0
7	Pneumonia	0	8	226	0
8	Atelectasis	0	80	154	0
9	Pneumothorax	0	8	226	0
10	Pleural Effusion	0	67	167	0
11	Pleural Other	0	1	233	0
12	Fracture	0	0	234	0
13	Support Devices	0	107	127	0

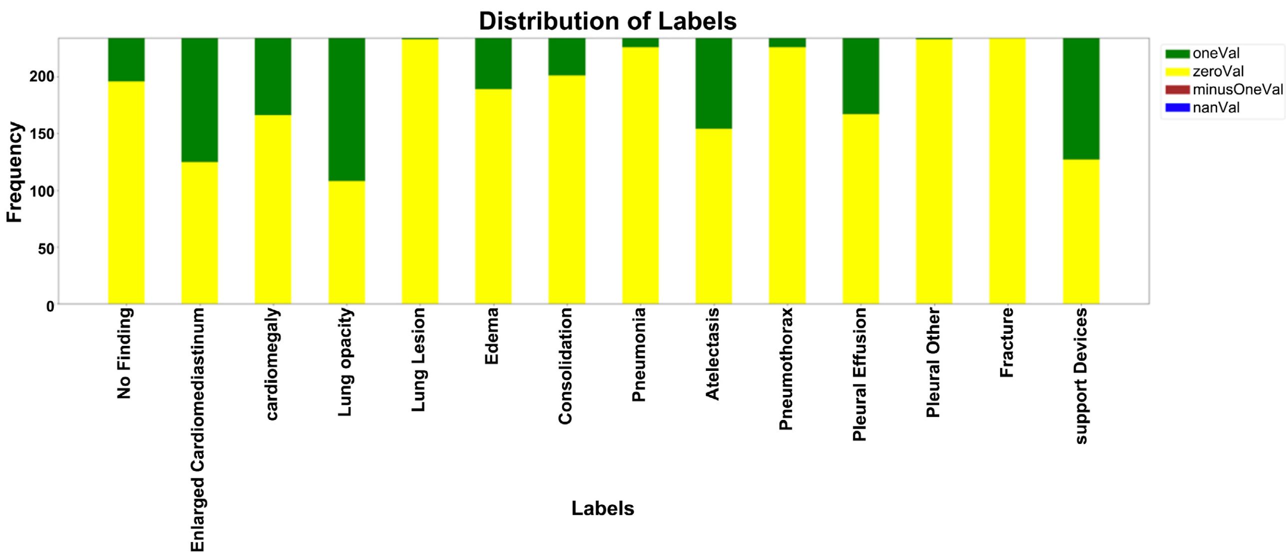


Figure 8. Distribution of test data and labels.

	Model Name	AUROC	Accuracy	F1 Score	Precision	Recall
0	CustomNet	0.753760	0.863197	0.245288	0.389082	0.200749
1	DenseNet121	0.783894	0.866112	0.278114	0.376906	0.240518
2	ResNet50	0.730819	0.845245	0.183340	0.282960	0.169079
3	Inception	0.692110	0.829782	0.116774	0.165283	0.096235
4	Vgg16	0.716072	0.838025	0.128222	0.186166	0.105025

Figure 9. Test metrics overall summary.

3.2. Test Metrics for Labels—AUROC

Densenet121 performed better in the case of individual labels also. Figure 10 shows AUROC values for various labels. Densenet121 model achieved AUROC ranging from 0.82 to 0.93 for the 5 labels those were part of the study conducted by Stanford ML group [1]. The performance was good for some other labels also like Pleural Other (AUROC: 0.97) and Lung Opacity (AUROC: 0.91). The worst performance was for Enlarged Cardiomediastinum (AUROC: 0.49).

3.3. Test Metrics for Labels—Accuracy

Densenet121 gave the best accuracy for individual labels. Figure 11 shows Accuracy values for various labels. Fracture, Lung Lesion, Pleural Other, Pneumonia and Pneumothorax achieved more than 95% accuracy. Enlarged Cardiomediastinum had the least accuracy of 53%.

					AUROC
Model Name	CustomNet	DenseNet121	Inception	ResNet50	Vgg16
Label					
Atelectasis	0.784740	0.824351	0.750568	0.753490	0.778571
Cardiomegaly	0.757884	0.834515	0.745836	0.724309	0.707123
Consolidation	0.841248	0.897332	0.672396	0.826624	0.817126
Edema	0.882775	0.883598	0.811875	0.838566	0.788948
Enlarged Cardiomediastinum	0.538789	0.485358	0.512073	0.568807	0.520954
Fracture	NaN	NaN	NaN	NaN	NaN
Lung Lesion	0.047210	0.227468	0.085837	0.098712	0.287554
Lung Opacity	0.864051	0.911229	0.782775	0.823927	0.821355
No Finding	0.865602	0.850161	0.843582	0.854189	0.872180
Pleural Effusion	0.865135	0.925552	0.780409	0.833676	0.833318
Pleural Other	0.935622	0.969957	0.849785	0.884120	0.656652
Pneumonia	0.586836	0.682522	0.373341	0.605642	0.481195
Pneumothorax	0.632190	0.680310	0.681416	0.611173	0.830752
Support Devices	0.854588	0.882258	0.705276	0.748915	0.715652

Figure 10. Test labels comparison—AUROC.

					Accuracy
Model Name	CustomNet	DenseNet121	Inception	ResNet50	Vgg16
Label					
Atelectasis	0.700855	0.752137	0.658120	0.675214	0.658120
Cardiomegaly	0.743590	0.713675	0.709402	0.709402	0.709402
Consolidation	0.858974	0.858974	0.858974	0.858974	0.858974
Edema	0.863248	0.863248	0.803419	0.811966	0.816239
Enlarged Cardiomediastinum	0.534188	0.534188	0.534188	0.534188	0.534188
Fracture	1.000000	1.000000	1.000000	1.000000	1.000000
Lung Lesion	0.995726	0.991453	0.995726	0.995726	0.995726
Lung Opacity	0.739316	0.782051	0.628205	0.756410	0.739316
No Finding	0.863248	0.807692	0.837607	0.858974	0.837607
Pleural Effusion	0.833333	0.871795	0.782051	0.773504	0.777778
Pleural Other	0.995726	0.995726	0.995726	0.995726	0.995726
Pneumonia	0.965812	0.965812	0.965812	0.965812	0.965812
Pneumothorax	0.965812	0.961538	0.965812	0.965812	0.965812
Support Devices	0.777778	0.769231	0.615385	0.679487	0.641026

Figure 11. Test labels comparison—Accuracy.

3.4. Confusion Matrix

Accuracy results may be misleading in some cases when the data set is unbalanced. We used confusion matrix to visualize true positives, false positives, true negatives, and false negatives as shown in Figure 12. Confusion Matrices for the best model, CustomNet and Densenet121 is given below.

3.5. CustomNet

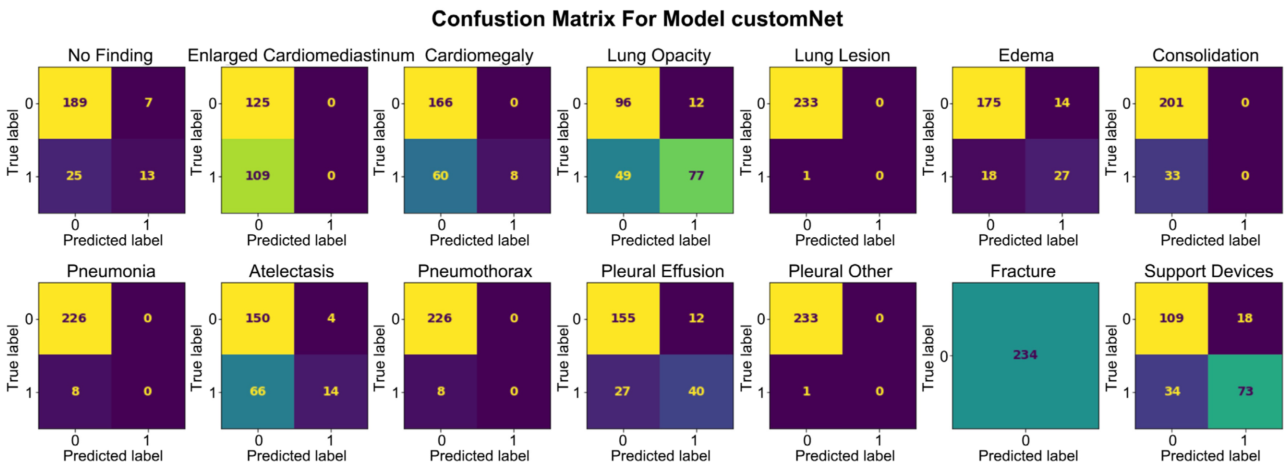


Figure 12. Confusion matrix for CustomNet.

3.6. DenseNet

Figure 13 demonstrates confusion matrix for DenseNet. The inability of the models to predict certain diseases are evident from the confusion matrix. For example, there were 68 positive cases for "Cardiomegaly" in the test data. Custom net predicted only 8 cases correctly where as DenseNet121 predicted 3 cases.

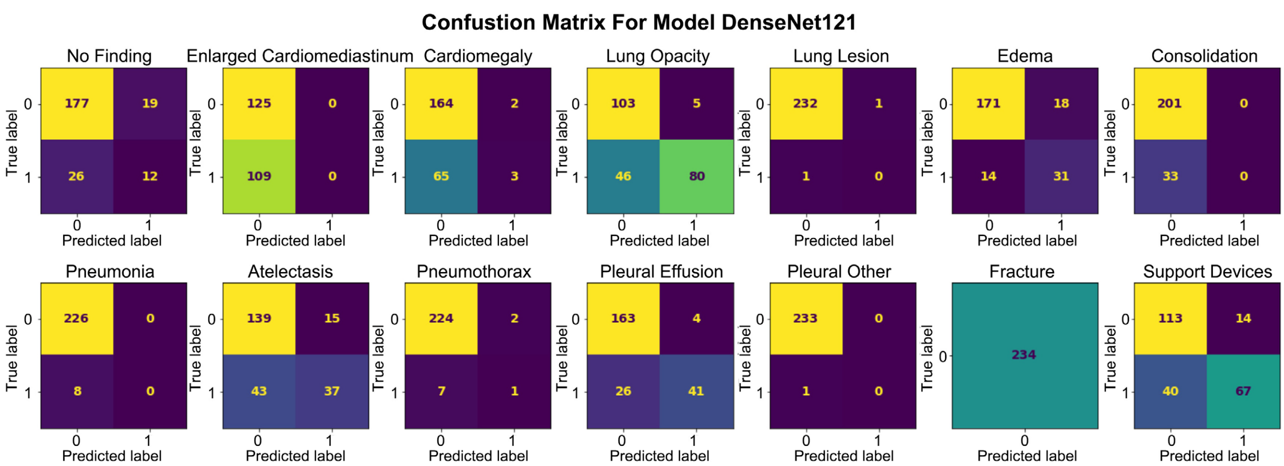


Figure 13. Confusion matrix for DenseNet.

4. Conclusion

The models were able to achieve ROC of about 0.78 and overall accuracy of about 87 percent. Dense121 pre-trained model gave the best performance for test data prediction. However, all models failed to accurately predict positive cases of certain diseases in spite of higher overall prediction accuracy rate. After examining the results, we hypothesize that these poor results are primarily due to lack of balanced training data. The number of positive cases in training data was very less compared to the negative class. Considering the uncertainty label as positive is not an effective approach for handling uncertainty in the dataset and is particularly ineffective on certain diseases. Overall, for 5 diseases, the best model was able to achieve AUROC comparable with baseline studies. Also model successfully predicted several other labels like Fracture, Lung Lesion, Pleural Other, Pneumonia and Pneumothorax with more than 95% accuracy.

Optimization

We realize that the class imbalance affected model's ability to predict positive cases of certain labels. For example, Cardiomediastinum has only 4% positive cases in train data and Pneumonia has only 3% positive cases. In the future, we wish to address this by experimenting with over sampling or under sampling techniques. The bias towards the dominant class can be reduced by altering the training data in order to decrease the imbalance.

Code Location

https://github.com/aravindsp/CS598_DL_Chest_X_RayClassification

Notebook

https://github.com/aravindsp/CS598_DL_Chest_X_RayClassification/blob/main/code/cs-598-multi-labelchest-x-ray-classification.ipynb

Python Script

https://github.com/aravindsp/CS598_DL_Chest_X_RayClassification/blob/main/code/cs-598-multi-labelchest-x-ray-classification.py

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Imaging and Radiology: MedlinePlus Medical Encyclopedia. https://medlineplus.gov/ency/article/007451.htm

[2] Deep Learning in Healthcare and Radiology. Aidoc Blog. https://www.aidoc.com/blog/deep-learning-in-healthcare

[3] Irvin, J., *et al.* (2019) CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison. https://www.aaai.org

[4] Nair, A., Nair, U., Negi, P., Shahane, P. and Mahajan, P. (2019) Detection of Dis-

eases on Chest X-Ray Using Deep Learning. http://cikitusi.com

[5] Palanisamy, N. and Santerre, J. (2019) Automated Pleural Effusion Detection on Chest X-Rays. *SMU Data Science Review*, **2**, Article No. 15. http://digitalrepository.smu.edu.Availableat https://scholar.smu.edu/datasciencereview/vol2/iss2/15

[6] Moradi, M., Madani, A., Karargyris, A. and Syeda-Mahmood, T.F. (2018) Chest X-Ray Generation and Data Augmentation for Cardiovascular Abnormality Classification. *Proceedings of SPIE*, Vol. 10574, 105741M. https://doi.org/10.1117/12.2293971

[7] Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M. and Summers, R.M. (2017) Chest X-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Honolulu, 21-26 July 2017, 2097-2106. https://doi.org/10.1109/CVPR.2017.369

[8] Johnson, A.E.W., *et al.* (2019) MIMIC-CXR-JPG, a Large Publicly Available Database of Labeled Chest Radiographs. http://arxiv.org/abs/1901.07042

[9] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. http://arxiv.org/abs/1409.1556

[10] Huang, G., Liu, Z., van der Maaten, L. and Weinberger, K.Q. (2016) Densely Connected Convolutional Networks. 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Honolulu, 21-26 July 2017, 2261-2269. http://arxiv.org/abs/1608.06993 https://doi.org/10.1109/CVPR.2017.243

[11] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, 27-30 June 2016, 770-778. http://image-net.org/challenges/LSVRC/2015 https://doi.org/10.1109/CVPR.2016.90

[12] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z. (2015) Rethinking the Inception Architecture for Computer Vision. 2016 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, 27-30 June 2016, 2818-2826. http://arxiv.org/abs/1512.00567 https://doi.org/10.1109/CVPR.2016.308

[13] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) ImageNet Classification with Deep Convolutional Neural Networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems*, Volume 1, 1097-1105. http://code.google.com/p/cuda-convnet

[14] Namdar, K., Haider, M.A. and Khalvati, F. (2021) A Modified AUC for Training Convolutional Neural Networks: Taking Confidence into Account. https://pypi.org/project/GenuineAI

Journal of Intelligent Learning Systems and Applications, 2022, 14, 57-70
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

Virtual Reality and Learning: A Case Study of Experiential Pedagogy in Art History

James Hutson, Trent Olsen

Department of Art History and Visual Culture, Lindenwood University, Saint Charles, MO, USA
Email: jhutson@lindenwood.edu, tolsen@lindenwood.edu

How to cite this paper: Hutson, J. and Olsen, T. (2022) Virtual Reality and Learning: A Case Study of Experiential Pedagogy in Art History. *Journal of Intelligent Learning Systems and Applications*, **14**, 57-70.
https://doi.org/10.4236/jilsa.2022.144005

Received: August 15, 2022
Accepted: October 30, 2022
Published: November 2, 2022

Abstract

While images are central to the discipline of art history, surprisingly little research has been conducted on the uses of digital environments for teaching in the discipline. Over the past decade, more studies have emerged considering the egalitarian space that can be used by students and teachers in web-based applications and social media. A body of literature has begun to emerge out of a small network of scholars and educators interested in digital humanities and art history, providing examples of how new tools can be integrated into the standard slideshow and lecture format of the field. At the same time, the latest technology that proves revolutionary for the field has had very little study-virtual reality (VR). Additionally, sensory evidence for digital art history and the creation of immersive interactive and multimodal environments for knowledge production is still underexplored. As multiple educational metaverses are currently under development, understanding best practices and pedagogical use of VR has never been timelier. This study seeks to review the pedagogical use of VR in art history current in the field and introduces results from a study of the most effective ways to use these immersive experiences using Bloom's revised taxonomy. Results confirm that the most effective method to structure VR assignments is to provide training on the technology, provide students with the necessary instructional material to introduce the concept, skill or technique to be learned, create or select an immersive experience that reinforces that topic, and conclude with a debrief or discussion about major takeaways from the experience.

Keywords

Virtual Reality, XR, Learning Outcomes, Art History, Experiential Pedagogy

1. Introduction

While many studies, especially in the digital humanities, have sought to integrate the latest technology and new media into the curriculum of art history, little has been published on the pedagogical application of virtual reality (VR) in particular [1]. For instance, in the collection of essays on Teaching Art History with New Technologies, new digital technologies and web-based applications for building concept maps and GIS are discussed, but VR was in its infancy [2]. Finch [3] confirms the surprising dearth of studies and points to the apparent hostility of digital environments in teaching in the discipline. A study by the Kress Foundation in 2011 confirms the assumption. The study was carried out with the Roy Rosenzweig Center for History and New Media at George Mason University and surveyed professionals on their use of digital art history [4]. The findings indicate an ambivalence or open hostility toward digital art history despite promising learning outcomes.

The most recent research into the use of VR in education has noted that the affordances include engaging learners in generative processing [5]. As with introducing a reading or a video, an immersive experience can either precede or follow a traditional lecture on a given topic in the field in order to provide greater insight or context [6]. Yet, little research has been undertaken to understand the pedagogical benefits of how instructors may leverage the immersive capabilities unique to the technology. Studies have thus far focused on qualities of presence, engagement, and immersion in general terms [7] [8] [9]. The manner and degree to which students are able to benefit from these qualities, however, depends upon their emotional engagement. Playing a video game or watching a video does seem engaging on their own, but in order to acquire emotional skills, technology employed needs to arouse emotional responses to support learning [10]. Certainly, cinema has the ability to evoke emotions in spectators, but does not induce emotions with a consistently high standard of reproducibility, prosocial change, and control [11]. Furthermore, there is the challenge of needing to induce emotions in an authentic way in order to simulate a user's real-life experience. Virtual reality is ideally suited to such experiences within virtual environments where stimuli can be controlled [12]. Once the experience has been identified and aligned with a learning outcome, positive results are noted and sense experience leveraged [13]. On the other hand, the relationship between the selected experience, outcome, and pedagogical methodology used in the classroom, especially the relationship between the use of VR and the instructional material, has yet to be thoroughly investigated.

The pedagogical approaches and methodologies best suited for the use of VR need be investigated. As such, this study seeks to determine the most effective pedagogical use of virtual reality using Bloom's revised taxonomy. Students from two sections of History of Western Art to 1300, which covers prehistory to the Renaissance, were instructed to complete the same virtual reality applications, which took them through a reconstruction of Pompeii. The first class had the experience before covering the material in class with a lecture and discussion. The second class had a lecture covering the material in class, and then had the VR experience without further debriefing. Student surveys, instructor feedback, and artifacts produced following the experience confirmed that introducing a topic to students that is reinforced with a VR experience and other supplementary materials then meeting together and discussing the topic in class yielded the best results. Furthermore, students claimed that with the limited interaction of the selected experience, VR is best suited for the lower levels of Bloom's revised taxonomy: "Explain ideas or concepts" (understanding), "Draw connections among ideas" (analyzing), and "Help recall facts and basic concepts" (remembering). The study confirms that the use of virtual reality improved student engagement and outcomes, as well as better understanding of topics covered.

2. Literature Review

With the ability to effectively transport students to any site in the world, including world culture heritage locations and museums, immersive realities are ideally suited to understand the context of a work. Through conveying critical changes in time, space or behavior through interaction and sensory immersion, these games can create thematic conceptual experiences that are themselves framed by contextual meaning. The field has already adopted emerging technologies to investigate cultural artifacts. Taking cues from the artworld itself, art historians disseminate reproductions of works first through printmaking and then through projections with magic lanterns. The modern age saw the rise of the slide carousel projector and now the ubiquitous ceiling-mounted LCD projector in classrooms across the world today. Art history has always sought out the most immersive methods to bring works to students. Furthermore, the ability of these emerging technologies to preserve, represent, and disseminate cultural heritage has received much attention in digital humanities scholarship [14] [15] [16] [17] [18]. But unlike the earlier technologies listed above, VR is not primarily a passive information delivery system. VR and gaming have the unprecedented educational ability to dynamically engage students and educators in a simulacrum. The three characteristics that act in concert to provide such an experience are outlined by Bekele and Champion [19] as the ability to: "1) establish a contextual relationship between users, virtual content, and cultural context, 2) allow collaboration between users, and 3) enable engagement with the cultural context in the virtual environments and the virtual environment itself." The features afford users the ability to engage with the experience, other users, and a deeper understanding of the context of the relationship between the three in a virtual environment.

Studies on the use of virtual reality (VR) in the art history classroom follow the increase in availability of hardware and relevant applications. Early studies focused on elements of immersion, presence, engagement, and educational potential of the new technology. For instance, a project by Casu, Spano, Sorrentino and Scateni [20] sought to leverage the lower cost of consumer hardware in developing an application for the teaching of Art History. Art Thief is another example of a game created at the California Institute of the Arts (2017) for CalArts

Game Makers Club. In this scenario-RPG, a young security guard named Olive must fend off an art thief in a museum, while simultaneously interacting with the museum staff and visitors to solve puzzles, etc. Unfortunately, the 2D game is limited in its interactions and is not as engaging as immersive content to be discussed [21]. Other examples allow for ease of use in configuring museums as desired. For example, ArtRift, a VR tool designed for art history students and teachers, allows for the configuration of virtual rooms in a museum with pre-selected artworks and is enhanced by multimodal annotation. Improving upon a traditional art history lecture where two works are compared and contrasted, the application allows for works to be juxtaposed with each other in each room and instructors add additional multimedia content, such as audio or textual descriptions. Among the immediate benefits of such virtual spaces, outlined by Casu, Spano, Sorrentino, and Scateni [20], is that comparisons can be made in a physical space that would never be possible in reality. Large sculptures, such as Michelangelo's David and Moses cannot ever be seen together as one is in Florence while the other Rome. In these simulated, virtual museum spaces, students can now compare the physical and stylistic elements of art as they were unable to do previously. In order to study the effectiveness of the application, students at Filippo Figari High School, Sassari were broken into two groups. The first was given access to works through a LCD projector on a wall, while the second through VR. At the close of the class, students were given the Instructional Material Motivation Survey instrument (IMMS) to assess their experiences. Students were queried on three areas: Attention Factor, Satisfaction, Relevance. All areas had qualitatively significant reporting and motivation was improved through the use of VR as opposed to traditional instructional methods.

As the previous example illustrates, previous research in the use of VR has primarily focused on secondary education [22]. As an example, Brownridge [23] outlined a curriculum to integrate VR into history and social studies classes in K-12 education. In the examples provided, students would take virtual field trips using Google Expedition (GE). A series of studies confirmed improved engagement and motivation once VR was integrated into coursework. The few studies that have been carried out in postsecondary education, however, have found the same positive correlations. For instance, Ghida [24] discusses how immersive reality has been used in his History of Western Architecture class, leveraging the ability to study a three-dimensional monument fully in three dimensions instead of an image or digital projection on a two-dimensional surface. Ghida provided students with specific monuments to view in Google Earth VR (released 2017) in order to experience monuments virtually in human scale. Given the advantages for architecture students, there is little surprise that the approach has since been adopted in architecture departments around the world, including Utah State University, MIT, Queensland University of Technology, Georgia State University, University of South California, the Chinese University of Hong Kong, Mount Saint Mary College, NY, and Florida State University.

In addition to providing the ability to tour monuments in three dimensions, research has also demonstrated that, as in the sciences, VR can be used to understand abstract concepts such as chronology in the field of art history, as well. Through a study of UK and Ukranian students at three levels from secondary to college, Korallo [25] sought to determine the effectiveness of using VLEs to assist in the teaching and learning of historical chronology in different fields. The study confirmed the difficulty in teaching the abstract nature of time for different learning levels. In order to learn the sequences of events and address this pedagogical issue, the groups were taught the sequences of events in a virtual environment. At the same time, control groups were shown the same events with texts and pictures, as well as with PowerPoint slides. Sets of parallel timelines were shown simultaneously, including music and art history, and the history of psychology, art and general history, respectively. The most beneficial experience was when undergraduate students were able to view three parallel timelines simultaneously within a continuous virtual environment. More specifically, using virtual environments assists in understanding historical events on a timeline and is a superior learning strategy than traditional techniques.

Another immersive experience was developed for a Renaissance art history class at the University of Indiana, Bloomington. Brennan [26] supported by the subject-matter expert, Dr. Giles Knox, developed four fresco cycles in the Unity game engine. With limitations of the HTV Vive headsets needing to be connected to a desktop computer, there were not enough units nor space to set up in the actual classroom, and thus the Virtual Reality Lab on campus hosted small groups of students outside of class time. After covering the material in class, students would then set up a time to explore the fresco that had just been covered, starting with the Scrovegni Chapel in Padua, Italy. 360-degree photographs of the cycles were imported as skyboxes into the game engine to support the build out. As prolonged movements often lead to VR sickness, students moved through the experience via teleportation between skyboxes/nodes. Smart history lectures were triggered when approaching the respective scenes. The production phase was followed by play testing and several iterations of the application to ensure the best user experience. In the study, students had the experience after covering the material in class. No data was collected regarding outcomes and taxonomic considerations.

The previous examples, while important early research into the field, dealt with limited pedagogical study of the use of virtual reality for art history coursework. In fact, few studies on VR pedagogy have been conducted in any area. The few that have do indicate a beneficial structuring of course activities to maximize the impact IR have on student engagement and interaction. For instance, Thorsteinsson [27] noted that in the Innovative Education (IE) study in Icelandic schools, Virtual Reality Learning Environment technology (VRLE) was used to support innovation. The technology was used to support online communications and collaboration between students and teachers to develop drawn solutions and descriptions of solutions to problems. The study by Thorsteinsson outlined the considerations for implementation and the various pedagogical spheres of learning that included the student's home, classroom, and VRLE environment. Students would generate content for the course for their homework, learn how to use the technology in the classroom, and ideate while using VR headsets and collaborate with other students. The steps set out for the curriculum included 1) finding needs; 2) brainstorming; 3) creating and choosing initial solutions; 4) concept drawing; 5) creating a description of the solution; and 6) presentation of solution. The sequence included introducing a concept/problem and training in the technology, then the use of VRLEs, and, finally, a debrief with a discussion and presentation. The steps paralleled those taken by the teacher over the course of instruction, which included 1) introduction; 2) basic training; 3) students reporting needs and problems; 4) brainstorming sessions with students; 5) students working in groups or as individuals to develop solutions inside of VRLEs; and 6) summarizing lessons for students.

While studies specific to the field of art history have focused on limited qualitative evidence, further investigation into when, how, and for what purpose VR should be leveraged in educational settings is necessary. For instance, given the limited availability of games, simulations, or experiences provided by applications between 2000 and 2015, studies naturally focused on the potential viability of the new technology as a tool to enhance learning. With the release of the Oculus Rift in 2012, HTV Vive in 2015, and Oculus Quest 2 in 2020, the barriers to adoption were largely removed. With greater access to the head-mounted displays (HMDs), educators and researchers could expand access to a wider demographic for a variety of studies. As studies now confirm, the positive correlation between the use of VR in education and learning outcomes and engagement, institutions of higher education are adopting the technology at greater rates. Recent polls also indicate that 90% of institutions will increase adoption of XR in the next five years, while only 12% are currently adopting XR broadly [28]. With educational metaverses in development, understanding the best practices and pedagogical strategies when implementing the technology has never been timelier.

3. Methodology

The mixed-methods study included data from surveys collected from students, instructor feedback and artifacts (short essays). The sample was collected from Lindenwood University, a private, four-year, liberal arts institution in the suburban ring of St. Louis, Missouri. Participants included 47 students from all four academic colleges from across the institution: Colleges of Education and Human Services, Arts and Humanities, Science, Health and Technology, and The Plaster College of Business and Entrepreneurship-enrolled in two General Education sections of History of Western Art to 1300. The two hybrid sections were offered in the Fall semester of 2021 and were designed to meet 50% of the allotted time face-to-face with additional activities, including readings, research, and recorded lectures, outside of class. The purpose of the project was to assess pedagogical best practices for the use of virtual reality (VR) through student perceptions, performance, and feedback coupled with instructor feedback and observations.

Students were tasked to engage with the Pompeii Experience available on the Oculus Rift. Additional learning activities focused on the city of Pompeii included a one-hour documentary, a brief online article, as well as classroom and recorded lectures and discussions. All students submitted a reflection paper in conjunction with the assignment. The experience occurred at the midpoint of the semester in addition to other standard course assignments that included a traditional formal analysis paper and two exams with slide identifications and comparison essays. Students in section one of the course were instructed to complete the VR experience, documentary, article, and online lecture before attending class. Following the experience, factual information regarding life and civilization in Pompeii was discussed in class. Students in the second section were instructed to view the experience after the content on Pompeii was presented in a classroom lecture.

This project utilized a mixed-methods approach to gather data, including qualitative (open-ended comments) and thematic (quantitative) results from an online survey. The survey instrument focused on the different methods for imparting information, as well as four different forms of media and thus informed the pedagogical considerations of VR. The survey was administered in Fall of 2021. Data collected afterwards gauged student demographics, feedback on the VR experience, asked for student preference for the order of introducing an art historical subject and supplementary learning materials, and how the technology would best be utilized in the six categories of Bloom's revised taxonomy. Students were then asked an open-ended question regarding their experience and what they felt VR was pedagogically best suited to accomplish. Students were contacted either through the University course management system or were emailed with links to online surveys. The survey was available for approximately two weeks at the end of the eight-week term and all data was collected using Qualtrics to ensure privacy and anonymity of responses. These results were sorted based on demographics (such as gender identity, major, age, etc.) and data were exported from the survey system. Descriptive statistics were calculated and used for comparisons between groups. The comparison of the two sections looked carefully at the level of engagement, discussion participation, excitement over the topic covered, understanding of material, and this was gauged via data gathered via student surveys, reflection papers, and outcomes on assignments and exams. The reflective essays students produced were evaluated along with the results of the surveys in order to glean more information on learning outcomes and more extensive feedback on the experiences.

4. Results

Of the 47 student respondents, 97.87% of participants were between 18 - 24 years of age; 61.7% identified as female, 36.17% male, and 2.13% non-binary;

76.60% identified as White, 14.89% Black or African American, 6.38% Hispanic or Latino, and 2.13% Asian. The majority (59.57%) of students were enrolled in the course to fulfill a General Education requirement, 36.17% enrolled to fulfill major requirements, and 4.26% as an elective. While most of the students (83%) reported being slightly familiar or more with VR technology prior to the class, 68.09% of them had also never used VR previously, while 27.66% reported using it on a monthly basis, and 4.26% of students reported using VR weekly Figure 1.

In terms of student perception of their preference for completing the VR activity prior to or after discussing the history of Pompeii in class, students preferred the chronology employed in their section. Of the students in section one who completed the VR and associated Pompeii activities before the activity debrief and lecture in class, 50% reported that have a VR experience prior to learning about Pompeii in class was better, while 20% felt it would be better after learning about the city in class Figure 2. Likewise, 50% of students in section two who completed the VR activity after discussing the history of Pompeii preferred this chronology, but none thought it would be better to complete the VR activity before class. In a follow-up question regarding when they considered it would be appropriate to conduct a VR experience, each group also were inclined to their experience, where 90% of the students section one felt it would be somewhat to extremely appropriate to complete a VR activity before coming to class, 70% of the students in section two found it somewhat to extremely appropriate to learn about Pompeii in class before completing their VR activity Figure 3.

Student perception of preparation to complete a VR experience also did not differ strongly enough to inform the chronology of the assignment. All respondents from section one felt somewhat to very prepared to complete the VR activity

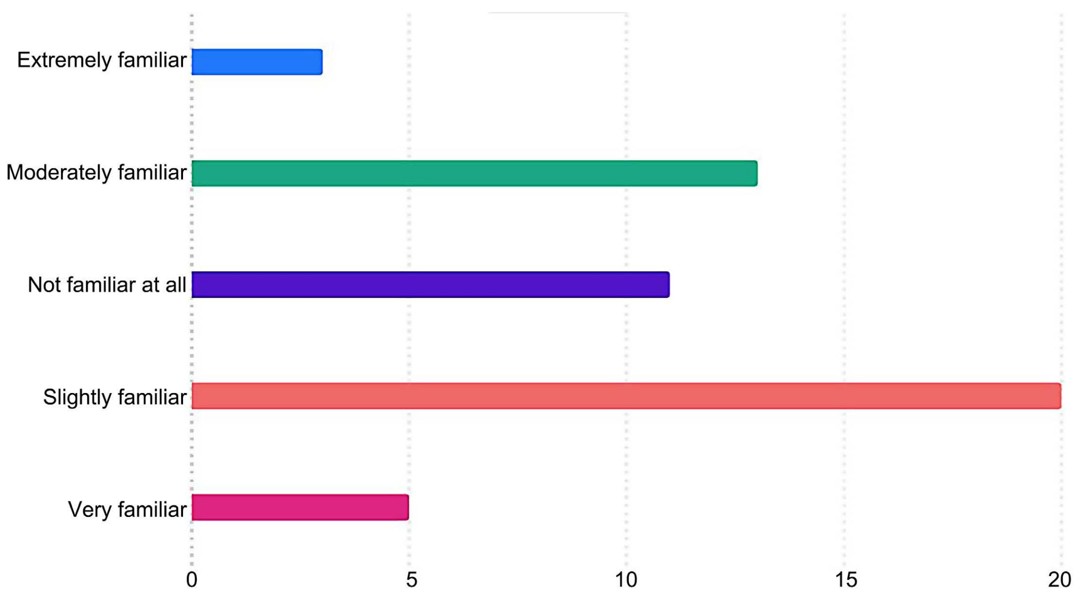


Figure 1. Student familiarity with virtual reality (VR) technology prior to class.

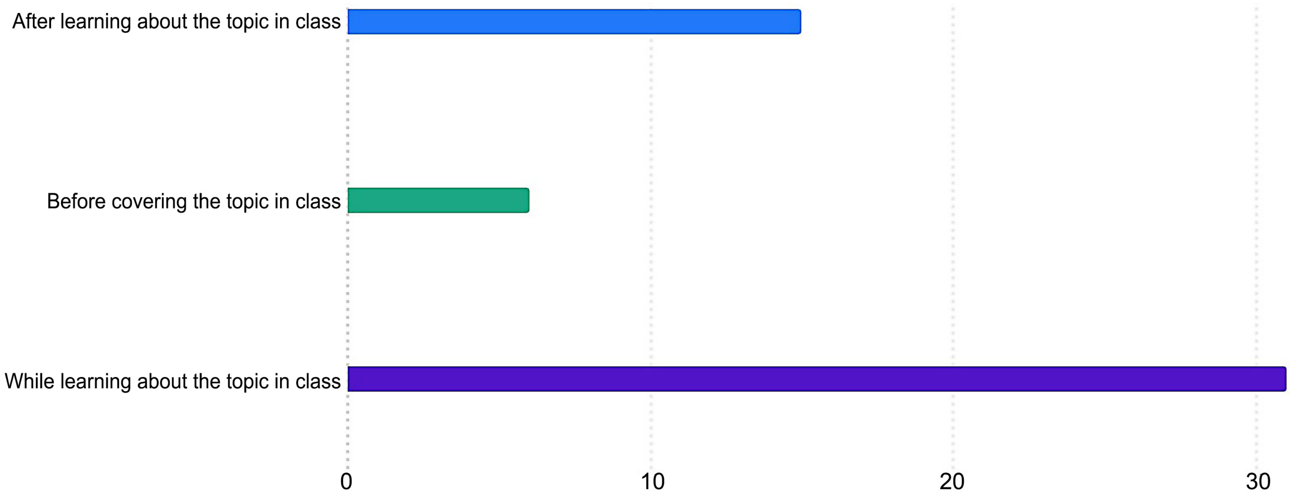


Figure 2. Student preferences for timing of VR experiences.

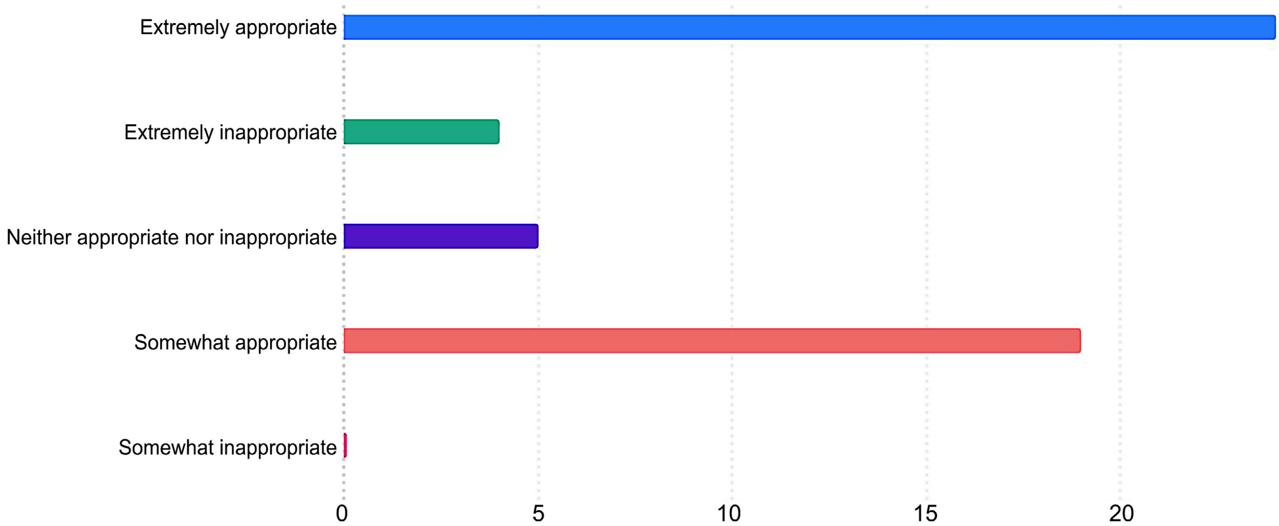


Figure 3. Student evaluation of appropriateness of timing of VR experiences.

before lecture. This rate actually decreased for section two, where 90% of students in this section felt somewhat to very prepared for the activity after learning about Pompeii. However, the students may have interpreted this sense of preparation to mean the instructions for how to complete the assignment, how to reserve the Oculus Rift headset, etc.

An area of consensus from students in both sections was their perception of the benefit of VR. 71.21% of students overall claimed that VR helped them learn better, with 25.53% selecting "maybe," for its helpfulness, and only 4.26% who responded "no". Beyond this, 91% of students found that learning about Pompeii through VR was more useful than reading about the topic, and 79.1% stated that it was more useful than watching a video. Therefore, students find VR the most useful, then watching a video, and, finally, reading about a topic. Several students use the open comment portion of their survey to leave comments such as: "It was super enjoyable especially because I've never learned like this! Sometimes I felt overwhelmed with what I needed to look at because there was so much, so I liked the VR experiences where it teleported yourself to different locations and explained the art in front of you."

Along these lines, students ranked the following categories adapted from Bloom's revised taxonomy. Based on the aggregate responses, respondents found the following to be the learning objectives best suited for VR (in order from most to least beneficial) Figure 4:

1) Explain ideas or concepts (understanding)
2) Draw connections among ideas (analyzing)
3) Help recall facts and basic concepts (remembering)
4) Use information in new situations (applying)
5) Justify a stand or decision (evaluating)
6) Producing a new/original work (creating)

From the instructor vantage point, changing the order of classroom lecture and student engagement with the Pompeii activities made a large impact on the shape of the class discussion. When introducing his students to Pompeii before they completed the additional assignments, Olsen reports that this order gives the instructor an open field to introduce students to the history and developments of Pompeii, with which they may not already be very familiar. Introducing new content in class can be an enjoyable component of teaching, especially when it is a topic that tends to pique student interest, such as the cataclysmic events at Pompeii. The clear disadvantage of this order is not being able to capitalize on the excitement students experienced engaging with the VR and additional learning materials. While they did record their observations in response papers, the level of enthusiasm wasn't transferred into the written response to the same degree as it was for students who discussed their experiences in class. And as the reflection paper was an individual assignment, it did not allow students to feed off of each other's experiences in discussing their observations. As a final positive element, when introducing new material to students before they

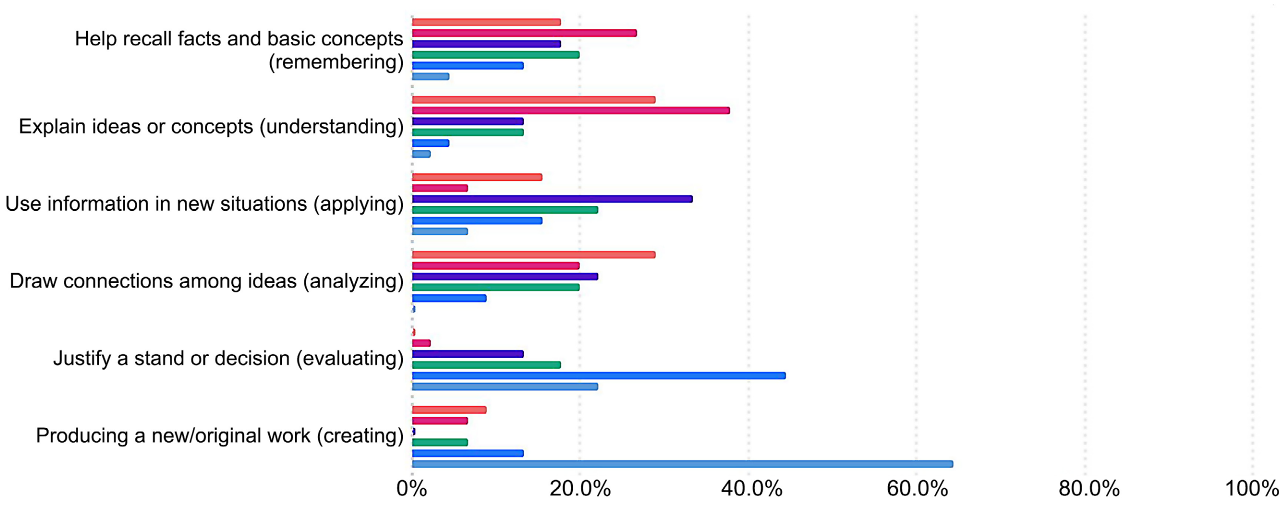


Figure 4. Student rank order of taxonomic benefits of VR experiences.

complete their activities as supplemental learning, one can verify that they have attained a level of proficiency and understanding. Completing additional materials after that will then provide even further clarification and insight to build on a base that they have already established.

In contrast, asking students to complete the VR assignment materials prior to discussing them in class carries much of the same benefits one encounters in instructing a flipped classroom. Students come to class with a solid level of understanding. But in this study, their experiential learning enhanced their understanding far greater than just engaging with a prerecorded lecture. Their understanding for the layout, remains, and daily operations of Pompeii was considerably more in-depth. That experience allowed classroom discussion to move beyond just quizzing or reviewing, and instead more time was used in analysis and application of learned concepts. What's more, students came to class excited to talk about their impressions. Olsen's sense is that part of this was due to the nature of discovering Pompeii and the charismatic and thought-provoking characteristic of the documentary his students watched. But these materials, in addition to the added layer of the novelty of VR technology and gaining a sense of experiential learning had a noticeable positive impact on his students. As students reported, the ability to see and move through the remains of Pompeii was exciting. However, they did report some minor frustrations with aspects with the VR application (glitchy movements, difficulty in navigation, VR sickness, less compatibility for students with glasses). One student articulated a clear point of benefit from this approach. He recounted his excitement when he had completed the VR experience before watching the documentary. When the video moved through an area of the city he had encountered in the application, he enthusiastically expressed "hey I know that area, I just 'walked' along that street!"

As reflected in the survey data, the majority of students expressed their preference for learning about Pompeii via VR over watching a video or reading an article. In class, several students iterated this preference, citing the imitation of corporeal engagement in the process of their learning as being a factor that was so appealing. They enjoyed the active nature of the application, that they could move through spaces in the city and view things of their choosing, which they stated was more generally more appealing than just viewing a video on a screen.

Recommendations

Students from each section expressed a preference for their own order of engagement with learning materials. What is clear, however, is that students from each section saw the use of VR as beneficial in their learning. The use of this application assisted in their learning, sufficiently targeting the areas of Bloom's taxonomy that involve "explaining ideas", "recalling facts", and "creating new connections" as the most beneficial aspect of this technology. Those that were prepared with the instructional material prior to the experience and had an active learning activity including a debriefing discussion thereafter performed better on assessments. And while an instructor can utilize any order of engagement activities and classroom discussion to positive results, the instructor of record for this course found that debriefing the activity afterwards promoted more stimulating discussion and engagement in class. One possible explanation would be that when students arrive in class with a deeper, preliminary understanding of a topic, as well as a general enthusiasm regarding the experiential learning that has been encountered, and then the experience can be leveraged for further analysis and more in-depth application during in-class activities.

5. Conclusion

The adoption and use of VR in postsecondary education is still in their infancy. More educational applications are being developed across disciplines for ready-made experiences to introduce, reinforce, or master content. As with other pedagogical practices, instructors should use best practices to ensure the most benefit can be gained from the use of this and other emerging technologies. Merely selecting an application with similar content or subject matter to a topic covered in class will not guarantee a positive correlation between the experience and student learning outcomes. Understanding the benefits and limitations of virtual reality will be crucial in the coming decade for higher education, and the first step is familiarizing oneself as an educator with the tools now available for deployment. One should also consider how such tools should be used in the classroom. As demonstrated with this study, and confirmed in another [27], successful use of VR technology in the classroom should consider Bloom's revised taxonomy and introduce the concept and problem, train students on the use of the technology, allow them to experience it, and then debrief on their experience, as well as reinforcing learning outcomes.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Luhmann, J. and Burghardt, M. (2022) Digital Humanities—A Discipline in Its Own Right? An Analysis of the Role and Position of Digital Humanities in the Academic Landscape. *Journal of the Association for Information Science and Technology*, **73**, 148-171. https://doi.org/10.1002/asi.24533

[2] Donahue-Wallace, K., La Follette, L. and Pappas, A. (2009) Teaching Art History with New Technologies: Reflections and Case Studies. Cambridge Scholars Publishing, Cambridge.

[3] Finch, J.A. (2015) Visual Literacy and Art History: Teaching Images and Objects in Digital Environments. In: *Essentials of Teaching and Integrating Visual and Media Literacy*, Springer, Cham, 251-264. https://doi.org/10.1007/978-3-319-05837-5_13

[4] Zorich, D. (2012) Transitioning to a Digital World: Art History, Its Research Centers, and Digital Scholarship. A Report to the Samuel H. Kress Foundation and the Roy Rosenweig Center for History and New Media, George Mason University, Fair-

fax.

[5] Mayer, R., Makransky, G. and Parong, J. (2022) The Promise and Pitfalls of Learning in Immersive Virtual Reality. *International Journal of Human-Computer Interaction*. https://doi.org/10.1080/10447318.2022.2108563

[6] Smith, D.K. (2014) iTube, YouTube, WeTube: Social Media Videos in Chemistry Education and Outreach. *Journal of Chemical Education*, **91**, 1594-1599. https://doi.org/10.1021/ed400715s

[7] Yanovich, E. and Ronen, O. (2015) The Use of Virtual Reality in Motor Learning: A Multiple Pilot Study Review. *Advances in Physical Education*, **5**, 188-193. https://doi.org/10.4236/ape.2015.53023

[8] Kavanagh, S., Luxton-Reilly, A., Wuensche, B. and Plimmer, B. (2017) A Systematic Review of Virtual Reality in Education. *Themes in Science and Technology Education*, **10**, 85-119.

[9] Rojas-Sánchez, M.A., Palos-Sánchez, P.R. and Folgado-Fernández, J.A. (2022) Systematic Literature Review and Bibliometric Analysis on Virtual Reality and Education. *Education and Information Technologies*, 1-38. https://doi.org/10.1007/s10639-022-11167-5

[10] Saleme, P., Dietrich, T., Pang, B. and Parkinson, J. (2021) Design of a Digital Game Intervention to Promote Socio-Emotional Skills and Prosocial Behavior in Children. *Multimodal Technologies and Interaction*, **5**, 58. https://doi.org/10.3390/mti5100058

[11] Aitamurto, T., Stevenson Won, A. and Zhou, S. (2021) Examining Virtual Reality for Pro-Social Attitude Change. *New Media & Society*, **23**, 2139-2143. https://doi.org/10.1177/1461444821993129

[12] Louie, A.K., Coverdale, J.H., Balon, R., Beresin, E.V., Brenner, A.M., Guerrero, A.P. and Roberts, L.W. (2018) Enhancing Empathy: A Role for Virtual Reality? *Academic Psychiatry*, **42**, 747-752. https://doi.org/10.1007/s40596-018-0995-2

[13] Morélot, S., Garrigou, A., Dedieu, J. and N'Kaoua, B. (2021) Virtual Reality for Fire Safety Training: Influence of Immersion and Sense of Presence on Conceptual and Procedural Acquisition. *Computers & Education*, **166**, Article ID: 104145. https://doi.org/10.1016/j.compedu.2021.104145

[14] Addison, A. and Gaiani, M. (2000) Virtualized Architectural Heritage: New Tools and Techniques. *IEEE MultiMedia*, **7**, 26-31. https://doi.org/10.1109/93.848422

[15] Papagiannakis, G., Geronikolakis, E., Pateraki, M., Bendicho, V.M., Tsioumas, M., Sylaiou, S., Thalmann, N., *et al.* (2018) Mixed Reality Gamified Presence and Storytelling for Virtual Museums. In: Lee, N., Ed., *Encyclopedia of Computer Graphics and Games*, Springer, Cham, 1-13. https://doi.org/10.1007/978-3-319-08234-9_249-1

[16] Adhani, N.I. and Rambli, D. (2012) A Survey of Mobile Augmented Reality Applications. The 1*st International Conference on Future Trends in Computing and Communication Technologies*, Malacca, 89-95.

[17] Anthes, C., García-Hernández, R.J.G., Wiedemann, M. and Kranzlmüller, D. (2016) State of the Art of Virtual Reality Technology. *The Aerospace Conference*, 2016 *IEEE*, Big Sky, 5-12 March 2016, 1-19. https://doi.org/10.1109/AERO.2016.7500674

[18] Bekele, M., Pierdicca, R., Frontoni, E., Malinverni, E. and Gain, J. (2018) A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage. *Journal on Computing and Cultural Heritage*, **11**, 7-36. https://doi.org/10.1145/3145534

[19] Bekele, M. and Champion, E. (2019) A Comparison of Immersive Realities and In-

teraction Methods: Cultural Learning in Virtual Heritage. *Frontiers in Robotics and AI*, **24**, Article No. 91. https://doi.org/10.3389/frobt.2019.00091

[20] Casu, A., Spano, L.D., Sorrentino, F. and Scateni, R. (2015) RiftArt: Bringing Masterpieces in the Classroom through Immersive Virtual Reality. In: Biasotti, S. and Tarini, M., Eds., *STAG*: *Smart Tools & Apps for Graphics*, The Eurographics Association, Eindhoven, 77-84.

[21] Champion, E. and Foka, A. (2020) Art History, Heritage Games, and Virtual Reality. In: Brown, K., Ed., *The Routledge Companion to Digital Humanities and Art History*, Routledge, London, 238-253. https://doi.org/10.4324/9780429505188-21

[22] Merchant, Z., Ernest, Cifuentes, L., Keeney-Kennicutt, W. and Davis, T. (2014) Effectiveness of Virtual Reality-Based Instruction on Students' Learning Outcomes in k-12 and Higher Education: A Metaanalysis. *Computers & Education*, **70**, 29-40. https://doi.org/10.1016/j.compedu.2013.07.033

[23] Brownridge, P. (2020) A Place in History: Virtual Reality in Secondary Education. Dissertation, Rowan University, Glassboro.

[24] Ghida, B. (2020) Augmented Reality and Virtual Reality: A 360° Immersion into Western History of Architecture. *International Journal of Emerging Trends in Engineering Research*, **8**, 6051-6055. https://doi.org/10.30534/ijeter/2020/187892020

[25] Korallo, L. (2010) Use of Virtual Reality Environments to Improve the Learning of Historical Chronology. PhD Thesis, Middlesex University, London.

[26] Brennan, M. (2018) Virtual Reality in the Humanities Classroom. University of Indiana, Bloomington.

[27] Thorsteinsson, G. (2013) Developing an Understanding of the Pedagogy of Using a Virtual Reality Learning Environment (VRLE) to Support Innovation Education. In: *The Routledge International Handbook of Innovation Education*, Routledge, London, 486-500.

[28] McCormack, M. (2021) EDUCAUSE QuickPoll Results: XR Technology. EDUCAUSE. https://er.educause.edu/articles/2021/12/educause-quickpoll-results-xr-technology?utm_source=Selligent&utm_medium=email&utm_campaign=tl_newsletter&utm_content=12-21-21&utm_term=_&m_i=21MZE5R0rcn85Y_9TXxmVDAqd8mHr5cmGEEsQbDyaVb-ZOl37UWIBsU5T4DOZKY3kFenipU86dSFFUNsRB40cOPAfNm9UaGM22Q&M_BT=54752865242

Journal of Intelligent Learning Systems and Applications, 2022, 14, 71-88
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

Application of Artificial Intelligence Algorithm in Image Processing for Cattle Disease Diagnosis

Bezawit Lake[1*], Fekade Getahun[2], Fitsum T. Teshome[3]

[1]Department of Computer Science, Dilla University, Dilla, Ethiopia
[2]Department of Computer Science, Addis Ababa University, Addis Ababa, Ethiopia
[3]Department of Agricultural and Biological Engineering, Tropical Research and Education Center, IFAS, University of Florida, Homestead, USA
Email: *mahiderlake@gmail.com

How to cite this paper: Lake, B., Getahun, F. and Teshome, F.T. (2022) Application of Artificial Intelligence Algorithm in Image Processing for Cattle Disease Diagnosis. *Journal of Intelligent Learning Systems and Applications*, **14**, 71-88.
https://doi.org/10.4236/jilsa.2022.144006

Received: May 31, 2022
Accepted: November 1, 2022
Published: November 4, 2022

Abstract

Livestock is a critical socioeconomic asset in developing countries such as Ethiopia, where the economy is significantly based on agriculture and animal husbandry. However, there is an enormous loss of livestock population, which undermines efforts to achieve food security and poverty reduction in the country. The primary reason for this challenge is the lack of a reliable and prompt diagnosis system that identifies livestock diseases in a timely manner. To address some of these issues, the integration of an expert system with deep learning image processing was proposed in this study. Due to the economic significance of cattle in Ethiopia, this study was only focused on cattle disease diagnosis. The cattle disease symptoms that were visible to the naked eye were collected by a cell phone camera. Symptoms that were identified by palpation were collected by text dialogue. The identification of the symptoms category was performed by the image analysis component using a convolutional neural network (CNN) algorithm. The algorithm classified the input symptoms with 95% accuracy. The final diagnosis conclusion was drawn by the reasoner component of the expert system by integrating image classification results, location, and text information obtained from the users. We developed a prototype system that incorporates the image classification algorithms and the reasoner component. The evaluation result of the developed system showed that the new diagnosis system could provide a rapid and effective diagnosis of cattle diseases.

Keywords

Deep Learning, Expert System, Livestock, Machine Learning, Neural Network

1. Introduction

Livestock is an important component of socio-economic development in Ethiopia. In terms of livestock population, the country ranks first in Africa and tenth in the world. Livestock contributes approximately 40% of the country's total agricultural output and 15% of the total gross domestic product [1]. However, various factors limit the potential economic benefits that could be obtained from livestock farming. One of the factors is the scarcity of veterinary practitioners that provide timely and accurate diagnoses and treatment. In Ethiopia, the veterinarian-to-animal ratio is 1:500,000 [2]. Another reason is the presence of endemic and transboundary diseases, which reduce the productivity of the cattle population [3]. Some of the factors that contribute to the spread of cattle diseases in the country include a lack of sufficient medical infrastructures, a lack of endemic and transboundary disease controlling protocol, a lack of local and central government attention, and the remoteness of livestock owners. A lack of priorly established endemic and transboundary disease controlling protocol could expose countries to a severe spread of animal diseases [4]. Disease outbreaks reduce the quality and productivity of animal products such as milk, skin, and hides, and this in turn results in trade restrictions on these products [5]. Cattle skin diseases are conditions that cause inflamed, irritated, or scaly skin, hair loss, changes in skin pigmentation, and visible growth [6]. Cattles suffer from a variety of skin problems, some of which are simple to treat while others are more difficult. Lumpy skin disease (LSD), bovine papillomatosis (warts), and dermatophytosis (ringworm) are common skin diseases in Ethiopia. Approximately 65% of skin and hide products are rejected due to poor quality caused by these diseases [7]. The prevalence of these diseases should be taken seriously to prevent them from negatively impacting the quality of finished products. As a result, quick detection and diagnosis of livestock diseases are critical for preventing the spread of outbreaks of livestock diseases. Currently, most medical diagnostic systems use either a text or image-based approach for acquiring symptoms. [3] developed a text-based expert system with hybrid reasoning consisting of the case and rule-based reasoning for the diagnosis of cattle diseases. The system accepts text-based symptoms from a user and then searches for a solution from a similar case in the case base. However, the accuracy of a symptom description of the conditions that were encountered using a text-based approach depends on the understanding of the person who describes the symptoms. Having a system that does not require the assistance of an expert to describe the symptoms will allow people without medical backgrounds to easily diagnose and treat the diseases. [8] developed a web-based expert system to diagnose infectious and non-infectious cattle diseases using rule-based reasoning. [9] developed an android application expert system to simplify disease detection and show brief information about cattle's using a rule-based forward reasoning engine. [10] developed a mobile-based diagnosis of skin diseases using case-based reasoning with image processing to detect diseases. They reported the system showed an encouraging performance with an accuracy of 83%. [11] developed a mobile application for the diagnosis of skin diseases using case-based reasoning with image processing to detect diseases by applying similar past problems. Thus, we need a system that can easily describe the symptoms of a disease in a way that is satisfactory for the correctness of the diagnosis. A hybrid diagnosis system has the advantage of representing symptoms more conveniently and easily compared to image or text-based systems. However, a hybrid diagnosis system is not sufficiently studied for cattle disease diagnosis. Therefore, the main objective of this project was to integrate artificial intelligence models such as convolutional neural networks and expert systems to develop a system that can diagnose cattle diseases easily and efficiently.

2. Materials and Methods

In this study, the integration of an expert system with deep learning image processing was proposed to develop a system that could quickly detect and diagnose cattle disease. The design of the cattle disease diagnosis system includes two modules *i.e.*, the expert system and image analysis modules.

Artificial intelligence (AI) techniques are known for their success in human and animal health studies [10]. An expert system (ES) is a computer program that can act as an expert in a specific field by analyzing and making decisions based on the knowledge base [12]. Digital image processing (DIP) is a technique that uses a computer system to manipulate images [13].

2.1. Hybrid Cattle Disease Diagnosis System

The design and development of a hybrid cattle disease diagnosis (HCDD) system consider the basic procedure used by veterinaries as well as three other considerations. The symptoms that were identified by visual inspection were brought to the system as an image. The images that were brought to the system were classified by the image analysis module to identify the symptoms category. In addition, the symptoms that were noted by palpation were brought to the system as text dialogue. The text interface was in form of a checkbox, where the user can select the appropriate symptoms. The location (GPS), where the image was captured, was extracted from the image to guide the diagnosis procedure. Smartphones tag all kinds of metadata to images. This metadata is known as exchangeable image file format (EXIF). Depending on the camera, EXIF data will store the current state of the camera when the photo was taken including date and time, shutter speeds, focal lengths, lens type, and location. The reasoning component reaches a final diagnosis result by integrating image classification results, location, and text information obtained from the users. The system architecture for the system is illustrated in **Figure 1**. The components of each module are described along with relevant techniques, algorithms, and considerations as the following.

The user interface (UI) module was responsible for communication between

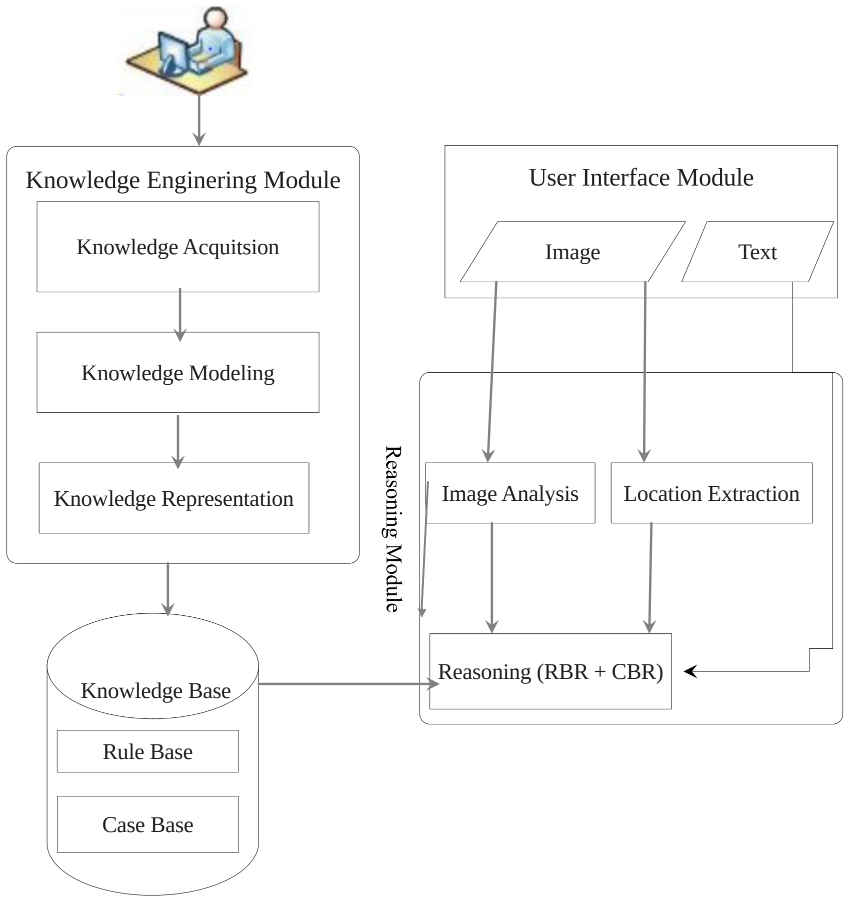


Figure 1. System architecture [RBR is rule-based reasoning and CBR is case-based reasoning].

the system and the user. It was used to collect important signs that occurred in symptomatic cattle. The UI has text and image-based information feeding components that allow users to feed critical information to the system.

The knowledge engineering module was responsible to build the required knowledge base for the proposed approach. It includes knowledge acquisition, knowledge modeling, and knowledge representation.

Knowledge acquisition was used to acquire knowledge and fact required for the diagnosis. The rules were extracted from the standard veterinary treatment guidelines such as Drug Administration and Control Authority (DACA) [14], black's veterinary dictionary [15] and Merck's veterinary manual [16]. The cases were collected from different case studies, and veterinary experts, and other literature were used as the main source of domain knowledge.

Knowledge modeling was responsible to model the acquired knowledge to understand the whole diagnosis and treatment techniques. It demonstrates the diagnosis procedures for skin diseases. The proposed disease diagnosis model considered all these features as shown in **Figure 2**.

Knowledge representation was a modeled knowledge that was represented in a way that was suitable for operation in a system. The case was represented using

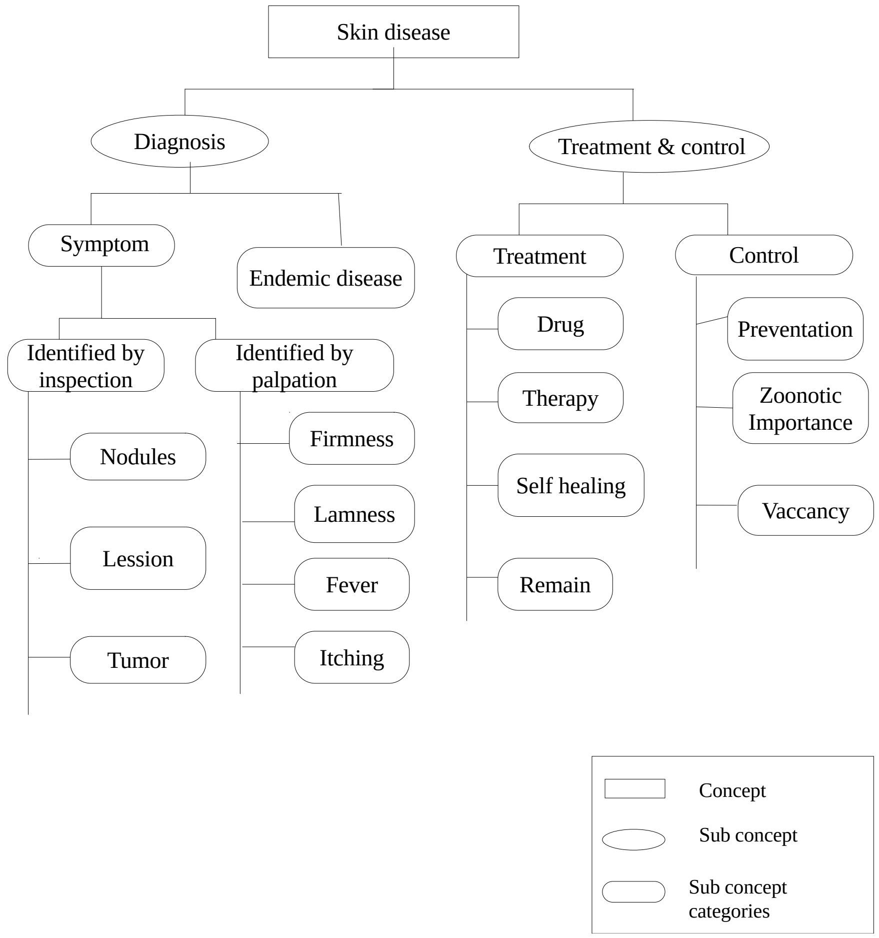


Figure 2. Disease diagnosis conceptual model.

cause-effect pairs and stored in the database. We used problem-solution case representation for location and their associated disease in that area. Data in the form of problem-solution pair was extracted for each case location and was stored in the database.

The design of the databases is shown in **Figure 3**. Table "CaseList" contained data about location information. The "LocationCaseDisease" table contained data about the diseases that occurred in each location. The "RuleMapping" table contained data about each rule and for what disease they were formulated. The rules acquired were represented using if … then representation, if a part contains the symptoms encountered and then part contains the consequence of the symptom. The treatment rules were represented using if … then representation if a part contains the disease and then part contains the treatment of the disease.

2.2. Image Analysis

The image analysis module was responsible for classifying the input image into different classes. In this study, a convolutional neural network classification algorithm was used to classify the images. A convolutional neural network (CNN) is a widely used image classification algorithm [17] [18] [19]. CNN automatically

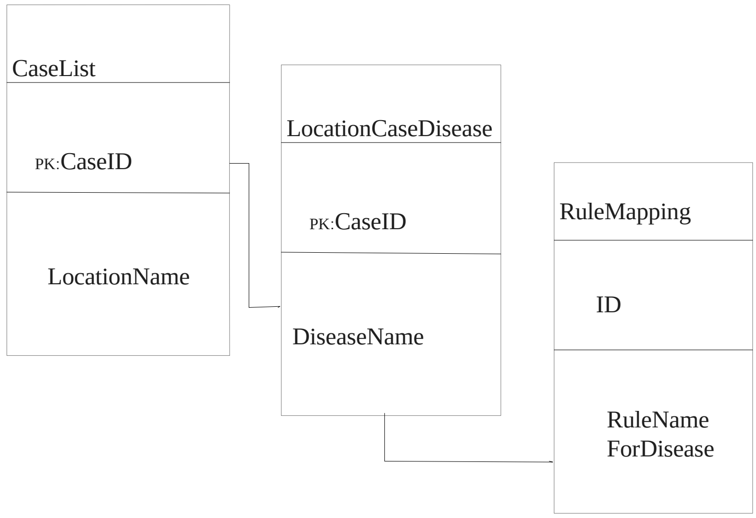


Figure 3. Database design [ID is identity document].

learns a hierarchy of features used for classification purposes. This is accomplished by successively convolving the input image with learned filters to build up a hierarchy of feature maps. Also, it is computationally efficient. The analysis starts with pre-processing the input images and stops when the classification result is found. Figure 4 shows the architecture of the image analysis module.

2.2.1. Pre-Processing

The pre-processing component was responsible to make the images suitable for the overall image analysis activity. The main tasks were resizing and normalizing. Resizing is one of the pre-processing techniques which brings the whole image to the same size. Since the images were collected from different sources and had different sizes, we decided on a new image size that best reflects the contents of the image with less processing time. All images in our dataset were resized into 200 × 200 pixels. Normalizing was another pre-processing technique we used before further processing. We normalized our data values down to a decimal between 0 and 1 by dividing the pixel values by 255.

2.2.2. Classification Model

The architecture of the model composes layers that were responsible for feature extraction and classifying the input image into one of the categories. In our model 200 × 200 RGB (red, green, and blue) image was passed through a stack of convolutional (Conv) layers, where we used filters with 3 × 3 (which is the smallest size to skim pattern). The convolution stride was fixed to 1 pixel. The spatial resolution was preserved after convolution. Spatial pooling was carried out by three max-pooling layers, which follow the conv layers. Max pooling was performed over a 2 × 2-pixel window, with stride 2. Figure 5 shows the proposed architecture of the CNN classification model.

In a convolutional network, the neurons are arranged in 3 dimensions width, height, and depth. The conv layers consist of a set of learnable filters. Every filter is arranged with width and height and extends through the full depth of the input

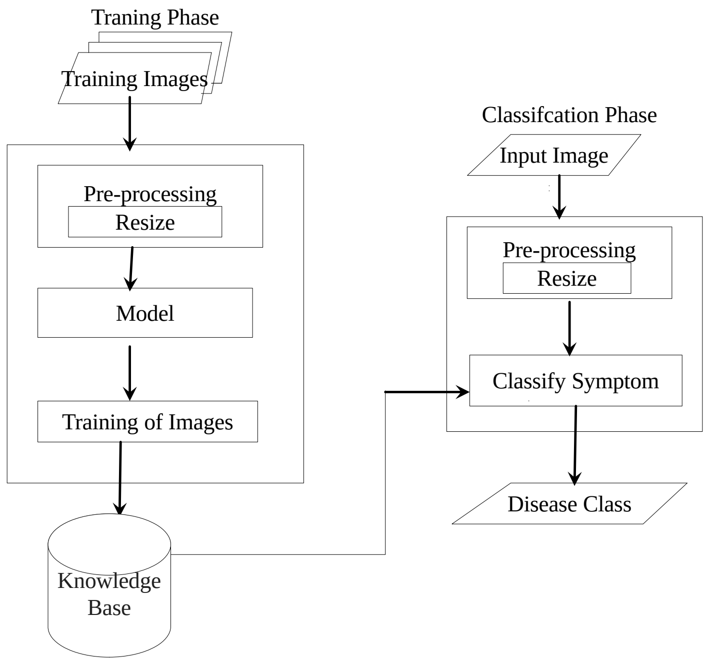


Figure 4. The architecture of the image analysis module.

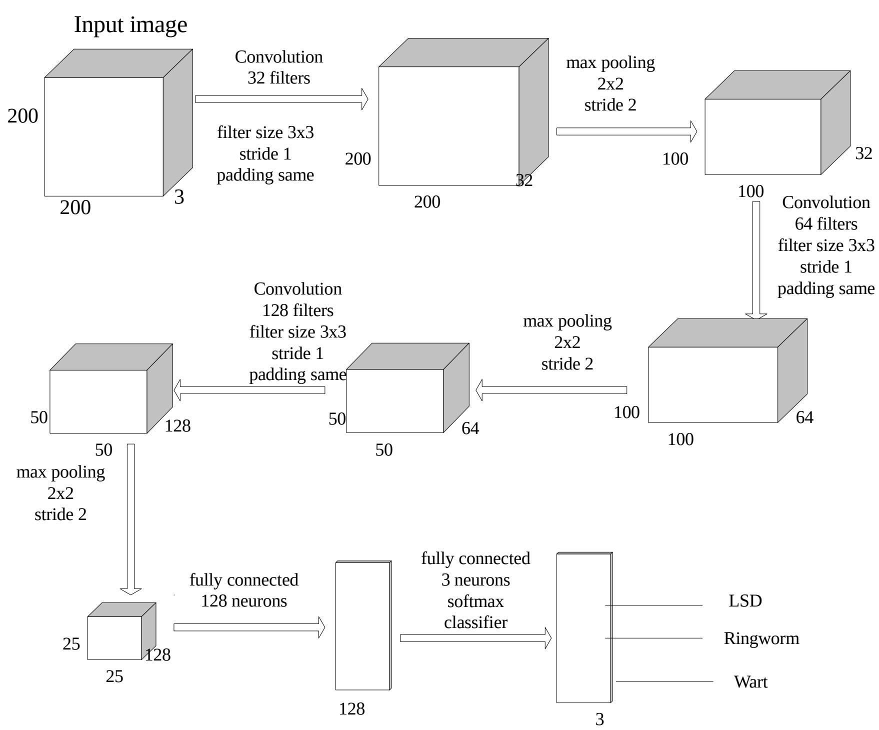


Figure 5. Classification model architecture [LSD is a Lumpy skin disease].

volume. The first conv layer accepted a 200 × 200 × 3 image and had a filter with 3 × 3 × 3 where the last 3 represented the depth of the filter which was like the input image. As we slide the filter over the width and height of the input volume, we produced a 2-dimensional activation map that gives the responses of that filter at every spatial position. Then we stacked these activation maps along the depth dimension and produced the output volume. In the first conv layer, 32 filters were applied which resulted in a 200 × 200 × 32 activation map. The max-pooling layer was performed a downsampling operation along the spatial dimensions (width, height), resulting in a volume of 100 × 100 × 32.

The second conv layer accepted the result of the first conv layer 100 × 100 × 32 and applied its 64 filters to the input which resulted in a 100 × 100 × 64 activation map. Max poling followed to perform downsampling and resulted in 50 × 50 × 64 volume. Its function is to progressively reduce the spatial size of the representation to reduce the number of parameters and computation in the network, and control overfitting. The third conv layer applied 128 filters. The final fully connected layer computed the class scores, resulting in a volume of size 1 × 1 × 3, where the 3 numbers correspond to a class score, among the 3 categories of our dataset. We observed that adding more layers did not improve the performance of our dataset. It leads to overfitting, increased memory consumption, and computation time. The removal of one layer from the model resulted in poor performance because the model had not generalized enough with a smaller number of layers in the model.

2.2.3. Training Classification Model

This component is where the model is trained for the task of image classification. The training starts by defining the optimizer, cost function, and metric. The workflow of the training is shown in **Figure 6**. The goal of a convolutional layer is feature extraction. When filter weights move over an image it checks for patterns in that section of the image. Filter weights change when the model is trained. In evaluation, these weights return high values if it thinks it is seeing a pattern it has seen before. The combinations of high weights from various filters let the network predict the content of an image.

The training was done using stochastic gradient descent with momentum (SGDM) optimizer, with a learning rate of 0.001, and a mini-batch size of 64 for 50 epochs. The training takes nine hours with HP Intel core i5-6200u CPU. The weight initialization was done using HE_UNIFORM initialization because it was compatible with the activation function, we use RELU for its less training time. The training (learning) in the model was presented as the function with the following form:

$$F(x) = f_5 f_4 f_3 f_2 f_1(x)$$

where, $f_1(x)$: Function learned on first conv layer, $f_2(x)$: Function learned on second conv layer, $f_3(x)$: Function learned on third conv layer, $f_4(x)$: Function learned on fourth hidden layer, and $f_5(x)$: Function learned on output layer.

The function value was computed by an iterative process of going and returning through layers of the model. The going was a forward propagation of the information, and the return was a backpropagation of the information. The first phase of forwarding propagation occurs when the network is exposed to the training data. Passing the input data through the network cause the neurons to

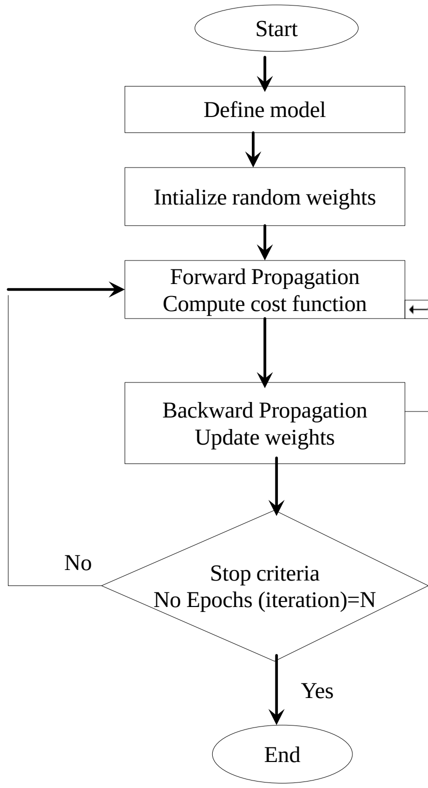


Figure 6. Workflow of the training algorithm [N is the number of iterations].

apply their transformation to the information they receive from the neurons of the previous layer and send it to the neurons of the next layer. When the data has crossed all the layers, the final layer was reaching a result of label prediction for those input data. After forward propagation loss function was calculated to estimate the error. The loss function was used to compare and measure how well or badly our prediction results compared to the correct result. Since our model task was the classification of the loss, we used Cross-Entropy Loss which can be defined by the following equation.

$$\text{CrossEntropyLoss} = -\left(T_y \log P_y\right) + \left(1 - T_y\right)\ \log\left(1 - P_y\right) \tag{1}$$

where T_y is the ground truth label for y and P_y is the predicted label for y.

The loss of information was calculated by propagating backward starting from the output layer to all neurons in the hidden layer. The neurons of the hidden layer only receive a fraction of the total loss, based on the contribution each neuron has contributed to the original output. This process was repeated, layer by layer, until all the neurons in the network have received a loss signal that describes their relative contribution to the total loss. This was done by calculating the partial derivate of the cost function relative to the weight of each neuron.

2.3. Reasoning Module

The reasoning module was responsible to seek information and relationship from the knowledge base to reach conclusion for the problem at hand. This component started with case-based reasoning (CBR) to acquire a disease that occurred in the input location case and rule-based reasoning (RBR) follows to reach a conclusion. The general high-level architecture of the reasoning module is shown in **Figure 7**.

The case which was like the input case was retrieved. Since the case base was a location case base, we needed an exact match when cases were retrieved. The algorithm used to retrieve cases was described in **Algorithm 1**.

Algorithm 1. Case retrieval from case base algorithm.

Input: input Case C_n, case stored in case base $C_m = [C_{m1}, C_{m2}, C_{mk}]$

Output: Similarity level

Begin:

For each case in the input case
Find the corresponding case in the stored case base
Compare the two values to each other
if $C_n = C_{mi}$
  $S_i = 1$
else $S_i = 0$
  END

Return Si

END

Adapt solution: the solution for the user query is decided if the retrieval stage identifies diseases, then the solution of the retrieved cases will be used. But if CBR retrieves no case, then the query is forwarded to the RBR component so that the final decision will be inferred from the rules. The workflow of the reasoning component was described in **Algorithm 2**.

Algorithm 2. Working algorithm of reasoning module.

Input: input Location case, classified image, text information

Output: Possible disease

Begin:

For input Location case
Location disease = result from CBR//list of diseases
If (Location disease = NULL)//Location case not found in KB
    Location = input Location case
    Possible disease = fire rules without mapping//no disease is stored on that location
    Store possible disease, Location in case base KB//as new case for the location
Else Fire rules with context//Location case found in KB
  Compute Mapping for Location disease with rule base
  Possible disease = fire rules with mapping
If (Possible disease = NULL)//no solution means it is new or epidemic disease for that Location
  Possible disease = fire rules without mapping
  Store possible disease, location in case base KB//as new case for the location

END

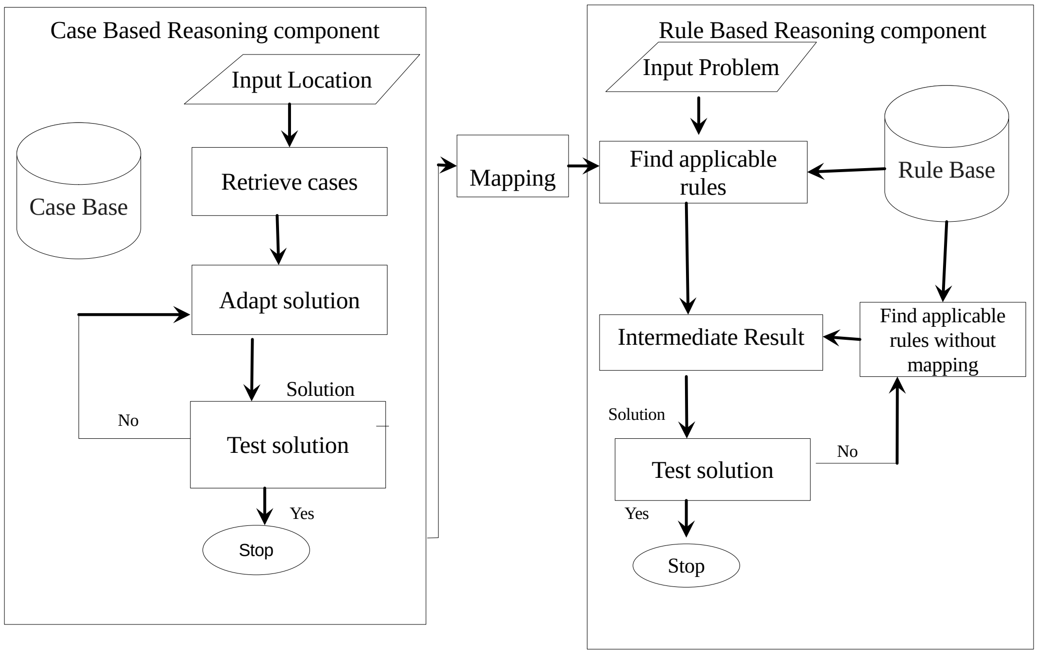


Figure 7. The architecture of the reasoner components.

2.4. Dataset Preparation

The process used for getting data ready for the classification model can be summarized in three steps: collect data, pre-process data and transform data. We followed this process iterative with many loops to prepare the dataset required.

Step 1. Collect data

This step involves collecting the available data needed to solve the problem. Image data were collected from Debre Markos University, Addis Ababa University-School of Veterinary Medicine, the Internet, and other secondary sources.

Step 2. Pre-process data

The pre-processing step was about getting the collected data into a form that can be easy to work. This step includes formatting and cleaning. The collected images were converted into JPEG format. JPEG format was selected because most of the collected images were in JPEG format and extraction of location information was possible with this format. We removed image data with unidentified labels so that we cannot get the proper label for them.

Step 3. Transform data

In the diagnosis system, we only needed the part of the cattle which was symptomatic. Therefore, we transformed the collected image by cropping the area where symptoms were presented. The original images are cropped into images containing symptoms regions. The cropping follows the following rules, including healthy and symptom parts, isolated symptoms taken individually, and widespread symptoms taken both as a whole and divided into regions. After transformation based on the specified rules, we transformed the input images into several images. The data we collected before and after transformation is shown in Table 1.

Table 1. Collected data before and after transformation.

No	Disease	Number of images collected	
		Before transformation	After transformation
1	Lumpy skin disease (LSD)	84	146
2	Ringworm	57	100
3	Wart	64	124

Our model required a large amount of labeled data. But getting enough data was a major problem in our cases which leads to the use of other techniques to expand our dataset. Data augmentation provides a means for increasing the quantity of training data available for machine learning and is particularly relevant when training deep learning systems from scratch [20]. Image augmentation is the process of taking images that are in a training dataset and manipulating them to create many altered versions of the image. It provides more images to train and expose our classifier to a wider variety of transformed images to make the classifier more robust. It has been widely used on small datasets for combatting over-fitting [19] [21]. Techniques of augmentation used in our dataset include horizontal and vertical flipping, zoom, shear, and rotation. **Figure 8** shows lumpy skin disease (LSD) infected cattle images after applying augmentation. After applying augmentation our dataset expands to 3990. Then we split into 90% training and 10% testing.

2.5. System Evaluation

The accuracy of the system is affected by the performance of the classification model, so evaluation was done on the model first. We evaluated the performance of the classification model using deep learning evaluation techniques accuracy and confusion matrix. Then, the accuracy of the entire system was evaluated with user evaluation. The performance of the model could be poor either due to overfitting or underfitting the data. The training of the model was plotted to see the possibility of overfitting and underfitting in the model.

The entire proposed system was evaluated by four people we selected randomly from different professions. Two of them were veterinarians, and the rest were individuals who have cattle. The selection assumed that those with a veterinary background can see and evaluate technical details while others may evaluate the applicability, accuracy, and importance of the system. Before starting the evaluation process, the system was explained in detail to the evaluators. The questionnaire we use for system evaluation is shown in **Table 2** where the user puts the weight for each evaluation. Weight values show the value of the evaluation, where 3 indicates the highest, 2 indicates the medium, and 1 indicates the lowest.

In addition, we constructed an evaluation matrix to evaluate the system. The evaluation matrix was a table with one row for each evaluator and columns that address evaluation questions shown in **Table 2**.

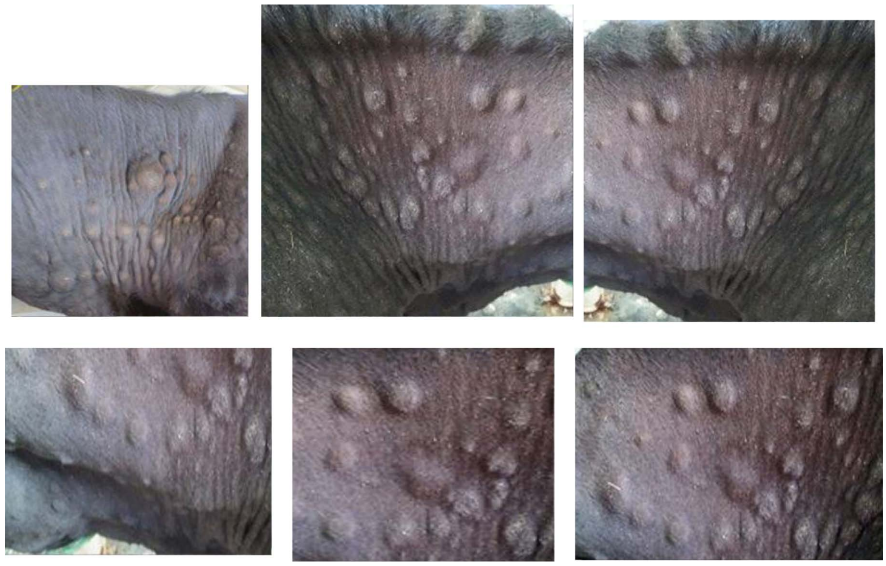


Figure 8. Images after augmentation techniques.

Table 2. System evaluation form.

No	Questions	1	2	3
1	Is the prototype system a user-friendly interface?			
2	How is the response time of the prototype system?			
3	Is the description of the symptom valid?			
4	Is the diagnosis result, correct?			
5	Is the treatment and recommendation, correct?			
6	How is the applicability of the prototype system?			
7	How is the performance of the prototype system?			

3. Results and Discussion

3.1. System Prototype

To demonstrate the validity of the proposed system, we developed a porotype system. The prototype shows the user interface and output of the system (**Figure 9**). The user interface allows the user to upload a photo and fill in the text information to the system. After the inputs are fed into the system, location extraction was done. Then, the image was passed to the processing components and the diagnosis result was given to the user.

3.2. System Evaluation

Overfitting happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. When the training accuracy is above the test accuracy it means the model is overfitting. Our model was not overfitting as shown in **Figure 10**. There was no significant difference between the value of training and test accuracy.

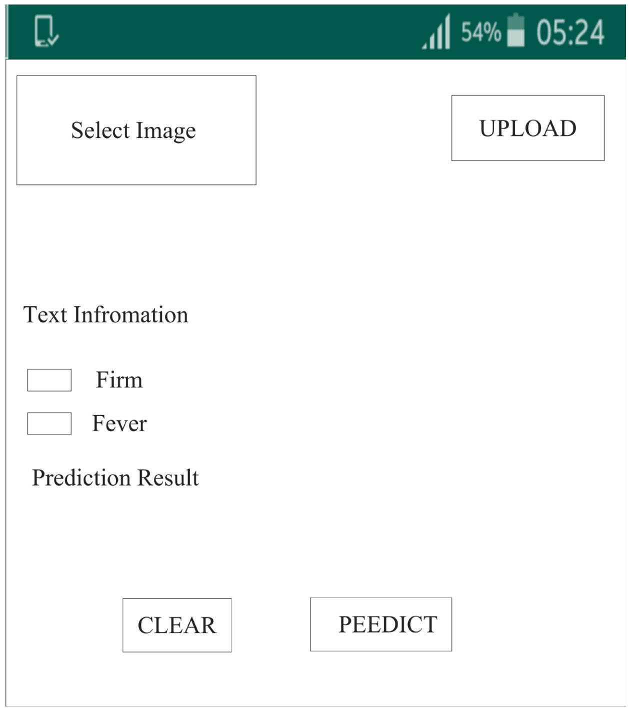


Figure 9. Prototype user interface.

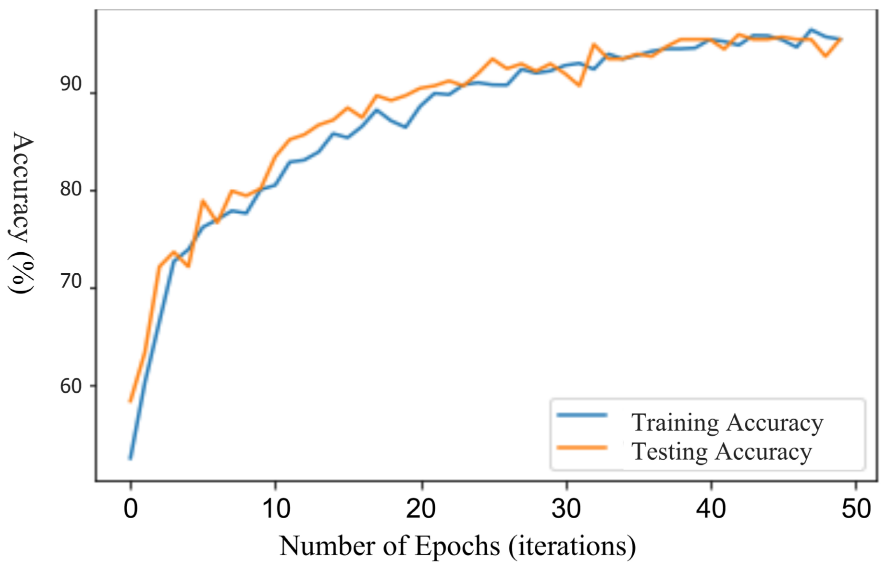


Figure 10. The plot of training and testing accuracy [X-axis is accuracy and Y-axis is number of epochs].

Underfitting refers to a model that can neither model the training data nor generalizes to new data. When validation loss is below the training loss the model is underfitting. As shown in **Figure 11**, our model was not underfitting. The model achieves 95% accuracy in 50 epochs.

To summarize the performance of the model we used a confusion matrix. In the confusion matrix, the number of correct and incorrect predictions was summarized with count values and broken down by each class. **Figure 12** shows

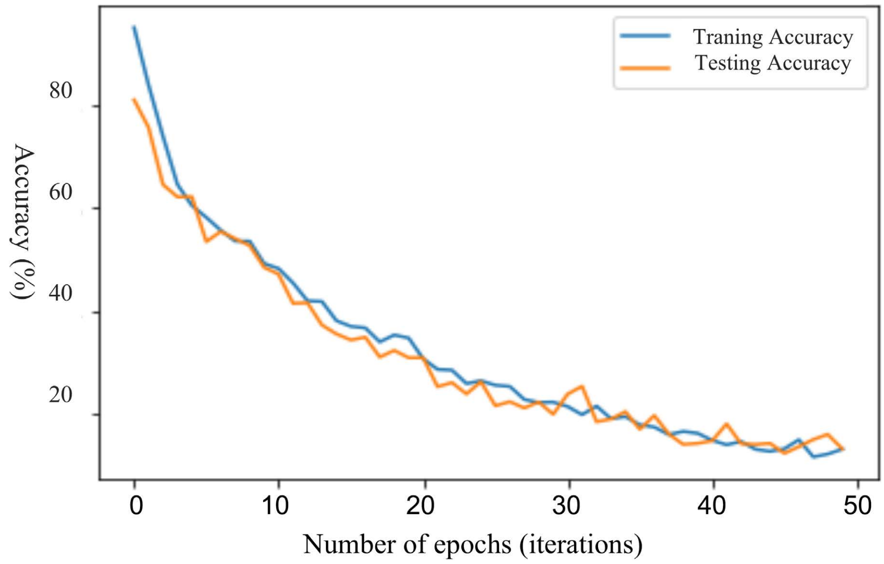


Figure 11. A plot of training and testing loss (X-axis is loss and Y-axis is number of epochs).

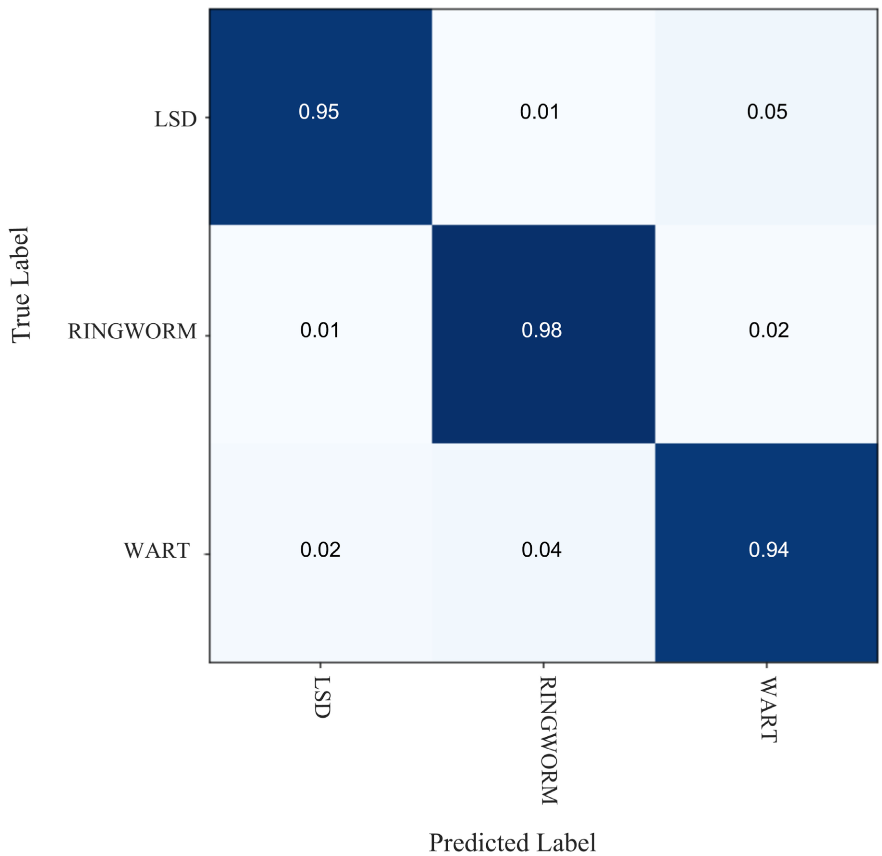


Figure 12. Confusion matrix of the model (LSD is a Lumpy skin disease).

how many of the images were misclassified and classified correctly. The ringworm classification was high compared to others because it has an easily distinctive feature from the other two classes. The misclassification of LSD as Wart and

Table 3. System evaluator's response.

No	Q1	Q2	Q3	Q4	Q5	Q6	Q7
1	3	3	3	3	3	2	3
2	3	3	3	3	3	2	2
3	3	2	3	3	2	3	3
4	3	3	3	3	2	3	3

Wart as LSD was noticed. Because LSD and Wart have more similar symptoms it is sometimes difficult to distinguish distinctive features among them.

The evaluator's response is recorded in Table 3. In Q1, the evaluation result shows that the prototype was user-friendly. It is easy for anyone to use because the interaction does not need any background knowledge. The evaluation result of Q2 shows that the response time for user queries was good. The response time is constrained by the storage and processing capability of the user's phone. For all users, the response time will be different. In Q3, the evaluation result shows a high value for the validity of symptom descriptions. So, representing symptoms using the image was a valid diagnosis method for cattle disease. In Q4, the diagnosis result shows a high value, which validates our diagnosis approach. In Q5, the correctness of treatment shows a lower result. The veterinarian recommends considering body mass index when a drug is prescribed, which our system did not include. In Q6 and Q7, the diagnosis result shows the performance and applicability of the proposed system were high. We can conclude that the integration of expert systems and image processing using deep learning gives an efficient and timely diagnosis of cattle diseases.

4. Conclusion

The potential economic benefit of livestock farming is impacted by numerous factors. One of these factors is the prevalence of livestock disease. This condition has been a significant factor that has affected the economic benefit of livestock farming. Quick detection and diagnosis of livestock disease are critical to prevent any outbreak of livestock disease from further spreading, as well as to improve the economic benefits that could be obtained from livestock production. In this study, the cattle disease diagnosis approach was developed by integrating an expert system and image processing using a deep learning algorithm. Diagnosis starts by acquiring information on the occurred disease through imageries and texts. Symptoms identified by inspection were acquired by capturing the image through mobile phones. Symptoms identified by palpation were presented to the system using text. To know the epidemic capability, location information was presented to the diagnosis system. Then, the image was pre-processed, and the class of the images was identified by the trained CNN model. The final diagnosis conclusion is drawn by the reasoner component of the expert system. The evaluation result shows that the developed approach effectively diagnoses cattle disease.

Acknowledgements

We would like to express our greatest gratitude to Dilla University for providing the funding for this research project.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Birhanu, T. (2014) Prevalence of the Major Infectious Animal Diseases Affecting Livestock Trade Industry in Ethiopia. *Journal of Biology, Agriculture and Healthcare*, **4**, 76-82.

[2] Tefera, M. (2010) Global Crisis and the Challenge of Veterinary Teaching in Ethiopia. *Global Veterinaria*, **5**, 294-301.

[3] Gebre-Amanuel, E.K., Taddesse, F.G. and Assalif, A.T. (2018) Web-Based Expert System for Diagnosis of Cattle Disease. In: *Proceedings of the* 10*th International Conference on Management of Digital EcoSystems*, *MEDES'*18, Association for Computing Machinery, New York, 66-73. https://doi.org/10.1145/3281375.3281400

[4] Domenech, J., Lubroth, J., Eddi, C., Martin, V. and Roger, F. (2006) Regional and International Approaches on Prevention and Control of Animal Transboundary and Emerging Diseases. *Annals of the New York Academy of Sciences*, **1081**, 90-107. https://doi.org/10.1196/annals.1373.010

[5] Donald, B., John, D. and David, H. (2006) USDA ERS—Economic Effects of Animal Diseases Linked to Trade Dependency [WWW Document]. https://www.ers.usda.gov/amber-waves/2006/april/economic-effects-of-animal-diseases-linked-to-trade-dependency

[6] Amin, D.M., Shehab, G., Emran, R., Hassanien, R.T., Alagmy, G.N., Hagag, N.M., Abd-El-Moniem, M.I.I., Habashi, A.R., Ibraheem, E.M. and Shahein, M.A. (2021) Diagnosis of Naturally Occurring Lumpy Skin Disease Virus Infection in Cattle Using Virological, Molecular, and Immunohistopathological Assays. *Veterinary World*, **14**, 2230-2237. https://doi.org/10.14202/vetworld.2021.2230-2237

[7] Belete, S. and Derso, S. (2015) Prevalence of Major Gastrointestinal Parasites of Horses in and around Mekelle (Quiha and Wukro). *World Journal of Animal Science Research*, **3**, 1-10.

[8] Lake, D. (2013) Web-Based Expert System for Cattle Diseases Diagnose. Masters, Addis Ababa University, Addis Ababa.

[9] Zamsuri, A., Syafitri, W. and Sadar, M. (2017) Web Based Cattle Disease Expert System Diagnosis with Forward Chaining Method. *IOP Conference Series: Earth and Environmental Science*, **97**, Article ID: 012046. https://doi.org/10.1088/1755-1315/97/1/012046

[10] Sallam, A. and Ba Alawi, A.E. (2019) Mobile-Based Intelligent Skin Diseases Diagnosis System. *The* 2019 *First International Conference of Intelligent Computing and Engineering* (*ICOICE*), Hadhramout, 15-16 December 2019, 1-6. https://doi.org/10.1109/ICOICE48418.2019.9035129

[11] Mieras, L.F., Taal, A.T., Post, E.B., Ndeve, A.G.Z. and van Hees, C.L.M. (2018) The Development of a Mobile Application to Support Peripheral Health Workers to Diagnose and Treat People with Skin Diseases in Resource-Poor Settings. *Tropical Medicine and Infectious Disease*, **3**, Article No. 102.

https://doi.org/10.3390/tropicalmed3030102

[12] Tripathi, K.P. (2011) A Review on Knowledge-based Expert System: Concept and Architecture. *IJCA Special Issue on Artificial Intelligence Techniques—Novel Approaches & Practical Applications*, No. 4, 21-25.

[13] Ravikumar, R. and Arulmozhi, D.V. (2019) Digital Image Processing—A Quick Review. *International Journal of Intelligent Computing and Technology* (*IJICT*), **2**, 11-19.

[14] DACA (2006) Standard Veterinary Treatment Guidelines (Drug Administration and Control Authority of Ethiopia). Veterinary Medicine. Infection. https://www.scribd.com/doc/48645004/STANDARD-VETERINARY-TREATMENT-GUIDELINES-Drug-Administration-and-Control-Authority-of-Ethiopia

[15] Boden, E. (2005) Black Veterinary Dictionary.

[16] Branion, H.D. (1963) The Merck Veterinary Manual. *Poultry Science*, **42**, 552-553. https://doi.org/10.3382/ps.0420552

[17] Alimboyong, C.R., Hernandez, A.A. and Medina, R.P. (2018) Classification of Plant Seedling Images Using Deep Learning. *The TENCON* 2018-2018 *IEEE Region* 10 *Conference*, Jeju, 28-31 October 2018, 1839-1844. https://doi.org/10.1109/TENCON.2018.8650178

[18] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. https://doi.org/10.1145/3065386

[19] Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A. and Baik, S.W. (2019) Multi-Grade Brain Tumor Classification Using Deep CNN with Extensive Data Augmentation. *Journal of Computational Science*, **30**, 174-182. https://doi.org/10.1016/j.jocs.2018.12.003

[20] Spyridonos, P., Papageorgiou, E.I., Groumpos, P.P. and Nikiforidis, G.N. (2006) Integration of Expert Knowledge and Image Analysis Techniques for Medical Diagnosis. In: Campilho, A. and Kamel, M., Eds., *Image Analysis and Recognition*, Springer, Berlin, 110-121. https://doi.org/10.1007/11867661_11

[21] Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J. and Johannes, A. (2019) Deep Convolutional Neural Networks for Mobile Capture Device-Based Crop Disease Classification in the Wild. *Computers and Electronics in Agriculture*, **161**, 280-290. https://doi.org/10.1016/j.compag.2018.04.002

Scientific Research Publishing

Journal of Intelligent Learning Systems and Applications, 2022, 14, 89-95
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

Digital Biomarker Identification for Parkinson's Disease Using a Game-Based Approach

Ilman Shazhaev[1], Dimitry Mihaylov[2,3], Abdulla Shafeeg[3]

[1]Management Department, Farcana, United Arab Emirates
[2]China Branch of BRICS Institute of Future Networks, Shenzhen, China
[3]Science Department, Farcana, United Arab Emirates
Email: abdulla.shafeeg@farcana.com

How to cite this paper: Shazhaev, I., Mihaylov, D. and Shafeeg, A. (2022) Digital Biomarker Identification for Parkinson's Disease Using a Game-Based Approach. *Journal of Intelligent Learning Systems and Applications*, **14**, 89-95.
https://doi.org/10.4236/jilsa.2022.144007

Received: September 11, 2022
Accepted: November 5, 2022
Published: November 8, 2022

Abstract

Despite the fact that their neurobiological processes and clinical criteria are well-established, early identification remains a significant hurdle to effective, disease-modifying therapy and prolonged life quality. Gaming on computers, gaming consoles, and mobile devices has become a popular pastime and provides valuable data from several sources. High-resolution data generated when users play commercial digital games includes information on play frequency as well as performance data that reflects low-level cognitive and motor processes. In this paper, we review some methods present in the literature that is used for identification of digital biomarkers for Parkinson's disease. We also present a machine learning method for early identification of problematic digital biomarkers for Parkinson's disease based on tapping activity from Farcana-Mini players. However, more data is required to reach a complete evaluation of this method. This data is being collected, with their consent, from players who play Farcana-Mini. Data analysis and a full assessment of this method will be presented in future work.

Keywords

Machine Learning, Biomarker, Parkinson's Disease

1. Introduction

Slowness of movement, or Bradykinesia, is a primary clinical symptom of Parkinson's disease (PD). Upper extremity bradykinesia may be evaluated using finger tapping tests, which are often used in neurophysiological exams. The gold standard in finger tapping assessment is the Movement Disorder Society Unified

Parkinson's Disease Rating Scale (MDS-UPDRS), which uses a 5-point rating scale to evaluate the condition. The MDS-UPDRS III's integer scale inhibits the identification of minor motor changes. Despite this, it is a complete evaluation. Inter-rater agreement is at best modest, with a score of 5. Because of this, there is an obvious need for techniques of evaluating motor dysfunction that are objective and consistent. An early and correct diagnosis is essential for patients to have access to the variety of available treatments and therapies. Early intervention may enhance life quality. The validation may aid experts in reducing differential diagnostic uncertainty, thereby eliminating the requirement for DaTSCAN and saving time and money. Priority may be given in triage to timely access to optimal therapy over wasteful recommendations of the worried-well.

It is possible to create unique games based on known behavioral correlates that are designed to put a player in a certain circumstance and track their actions. Custom assessment games have been designed to measure several elements of physical [1] and mental [2] well-being. The activity traces that are left behind by natural interactions with digital games may be employed as a digital biomarker of health and health deterioration.

As a result, in this paper, we hypothesize that changes in the patterns of finger movement while playing the Farcana-Mini game might be utilized to differentiate and categorize people with Parkinson's disease from those without the condition in the early stages of the disease. The difficulties in identifying PD stem from the lack of a conclusive test; at present, the condition must be diagnosed based only on clinical and observational criteria. Numerous symptoms of PD are unclear and are shared with other neurodegenerative and non-neurodegenerative disorders. The Unified Parkinson's Disease Rating Scale (UPDRS) is a tool for evaluating Parkinson's disease that is based on a score determined from a physician's neurological assessment; hence, it is a subjective measure that lacks objectivity, reproducibility, and sensitivity [3].

The majority of games need motor input to play. Touch input (*i.e.*, in mobile games); mouse and keyboard input (desktop games); and controller input, which comprises of buttons to push and thumbsticks or small joysticks to manipulate, are the most common types of input devices used in games (*i.e.*, in console games). Occasionally, gaming consoles have cameras that record the user's motions (e.g., Microsoft Kinect and Sony PlayStation Camera). Varying degrees of motor coordination are required to play different games: Many games demand complicated sequences of input (e.g., Street Fighter), whilst others require a relatively basic motor action, but require it to be performed rapidly and often (e.g., Cookie Clicker) or in conjunction with cognitive decisions (e.g., Farcana-Mini) [3].

2. Literature Review

Decho Surangsrirat *et al.* [4] have used the mPower dataset to examine how finger tapping activities on a mobile phone correlated with the MDS-UPDRS I-II and PDQ-8. mPower is mobile application-based research used for the study of the progression and diagnosis of Parkinson's disease. Today, it is the biggest open-access, mobile Parkinson's Disease study. Any public researcher was given access to data from seven modules with 8320 individuals who gave data for at least one task. Demographics, MDS-UPDRS I-II, PDQ-8, memory, tapping, voice, and walking are all included in this package. It is one of the activities that is straightforward to complete and has been evaluated quantitatively for PD measurement. They have only included individuals who completed both the tapping activity and the MDS-UPDRS I-II rating scale. According to the results of the statistical analysis, subjects were divided into three severity groups using the tapping characteristics. Based on the MDS-UPDRS I-II and PDQ-8 scores, each group indicated a distinct degree of Parkinson's disease (PD) severity.

Using 75,048 tapping accelerometer and position data, Kaiwan Deng *et al.* [5] built deep learning algorithms that generated an AUC (Area under the Receiver Operator Characteristic Curve) of 0.933 when detecting PD patients among 6418 individuals. The performance of tapping was superior to that of gait/rest and voice-based models derived from the same population used as a benchmark. When the three models were combined, the AUC increased to 0.944. Notably, the models not only corresponded significantly with patient-reported symptom levels, but also outperformed them in identifying PD. This work illustrates the complementary predictive potential of tapping, gait/rest, and speech data, and creates models based on integrative deep learning for detecting PD.

Noreen Akram *et al.* [6] developed a novel distal upper-limb exam called Distal Finger Tapping (DFT). Kinetic metrics include kinesia score (key taps over 20 s), akinesia time (mean dwell-time on each key), and incoordination score (IS20, variance of travelling time between key taps). To design and analyze a keyboard-tapping test for PD distal motor function. 55 PD patients and 65 controls took DFT and BRAIN. The MDS-UPDRS-III finger tapping sub-scores were correlated with test results. Nine more PD patients were recruited for motor monitoring. KS20 performed best, with 79 percent sensitivity and 85 percent specificity; AUC = 0.90. DFT and BRAIN enhanced discrimination (AUC = 0.95) KS20 correlated moderately with the MDS-UPDRS finger-tapping sub-score (Pearson's $r = 0.40$, $p = 0.002$). DFT They observed minor variations in motor fluctuation states that were not reflected in MDS-UPDRS-III finger tapping sub-scores. The DFT test assesses distal movements in PD and may be used to evaluate motor problems longitudinally. Lopez-de-Ipina Karmele *et al.* [7] developed a unique technique based on the integration of handwriting and neuroimaging data for the early clinical identification and monitoring of ET. A computerized Archimedes' spiral exercise measures fine motor abilities and correlates with ET MRI biomarkers. With their novel modeling technique, they offered a supplementary and promising tool for the clinical diagnosis of ET as well as tremors from a wide variety of sources.

3. Proposed Method

Farcana-Mini as shown in **Figure 1**, is a 2D game. As the character in the game

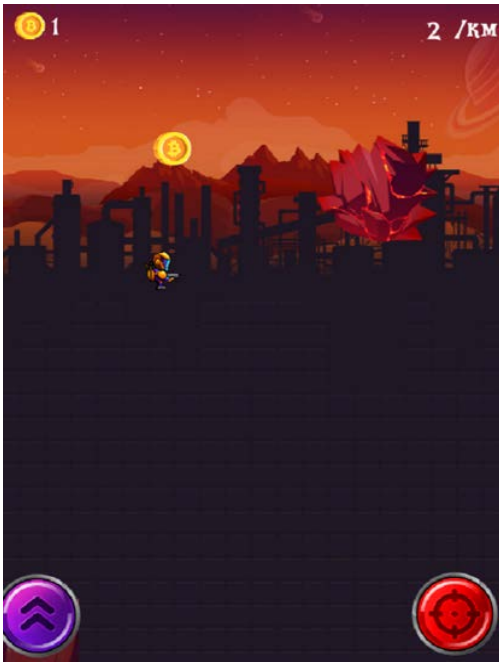


Figure 1. Farcana-mini.

moves forward player is required to continuously keep tapping the purple button at the lower left of display to avoid obstacles along the way as shown in Figure 1. The red button on the lower right of display can be used by player to shoot bitcoins at enemies or obstacles. If the character gets in contact with an obstacle or enemy the game restarts.

3.1. Data Collection

The Farcana-Mini is currently being played by thousands of players around the world and their finger tapping activity is being recorded and stored in a cloud server. With the consent of the players, we use this data for research. The assumption is that Parkinson's disease (PD) patients would have worse coordination and slower movements during gameplay compared to healthy individuals. This will result in erroneous and/or shifting tapping contact point locations.

In today's web2 ecosystem, service providers have complete control over the obtained user data. While the initial intended use of such data is largely for smart IoT systems and device control, the data is frequently utilized for additional uses that the users have not explicitly approved to. To preserve the integrity of data acquired by Farcana-mini players, we deploy Nevermined's innovative data exchange and storage structure (nevermined.io). It promises to provide users with complete data privacy control by seamlessly combining smart contracts, Data in Situ Computation, Federated Learning, and Provenance-based data integrity testing and verification in cloud settings for the usage of Service Execution Agreements (SEAs) between parties. The data owner not only controls who can have what access to his/her data, but also be ensured that the data is used only for the intended purposes.

3.2. Prediction Model

Neighborhood Components Analysis (NCA) is a distance metric learning algorithm that aims to improve nearest neighbor classification accuracy over standard Euclidean distance. The algorithm directly maximizes the one-to-one elimination (KNN) stochastic k-nearest neighbor estimator on the training sample. It can also learn a low-dimensional linear projection of the data, which can be used for data visualization and quick classification.

We focus on stochastic KNN classification. The thickness of the connection between a sample and another point is proportional to their distance and can be thought of as the relative weight (or probability) that a stochastic nearest neighbor prediction rule would assign to that point. In the original space, the sample has many stochastic neighbors from different classes, so the correct class is unlikely. However, in the predictive space learned by NCA, the only stochastic neighbors with a non-negligible weight are in the same class as the sample, which guarantees that the latter will be well classified.

NCA can be used for controlled dimensionality reduction. The input data is projected onto a linear subspace consisting of directions that minimize the NCA goal. The desired dimension can be set using the *n*_components parameter. For example, the following figure shows a comparison of dimensionality reduction using Principal Component Analysis (PCA), Linear Discriminant Analysis (Linear Discriminant Analysis), and Neighborhood Components Analysis on a Digits dataset, a dataset with size where the dataset is split into training and test sets of equal size and then standardized. For evaluation, the classification accuracy of 3 nearest neighbors is computed from the 2-dimensional predictive points found by each method. Each data sample belongs to one of 10 classes.

The goal of NCA is to learn the optimal linear transformation matrix of size, which maximizes the sum over all samples (*n*_components, *n*_features) i of the probability p_i that i is correctly classified, that is:

$$\operatorname{argmax}_L \sum_{i=0}^{N-1} p_i \tag{1}$$

involving $N = n$_samples and pi, the probability of sample i being correctly classified according to the stochastic nearest neighbor rule in the learned nested space:

$$p_i = \sum_{j \in C_i} p_{ij} \tag{2}$$

where C_i—is a set of points of the same class as sample i, and p_{ij} is softmax at Euclidean distances in the nested space:

$$p_{ij} = \frac{\exp\left(-\left\|Lx_i - Lx_j\right\|^2\right)}{\sum_{k \neq i} \exp\left(-\left\|Lx_i - Lx_j\right\|^2\right)} \tag{3}$$

3.3. Machine Learning Feature Selection

During data processing, distinguishing features between individuals who have signs of PD and individuals with no signs of PD are selected. These features include: Signal peak; Power spectral density average; Deviation standard of the power spectrum. In order to verify our work, we will use 10-fold cross validation to assess the predictive models' accuracy. In 10-fold cross validation method there are 10 iterations of cross validation in which a total of 90% of the entire training set is chosen at random, and 10% is utilized as a holdout group for testing and validation. In this study, there are two data groups: the training group and the test group. In the training group, the determinant attributes (processed data) and the class corresponding to each individual was provided; in the test group, only the attributes were provided, and the classification algorithm determined which group each individual belonged to. We will use the K-NN Classifier algorithm [8].

4. Conclusion and Future Work

In this paper, we have reviewed some of the methods found in the literature that were developed to identify problematic digital biomarkers. Most of the methods have performed quite well, but to the best of our knowledge, this is the first research to identify problematic digital biomarkers from data gathered while playing commercially available games that are meant for entertainment. We proposed a machine learning approach for early identification of hand tremor which is considered a sign of Parkinson's disease. We require more data to reach a complete evaluation. This data is being collected, with their consent, from players who play Farcana-Mini. Data analysis and a full assessment of this method will be presented in future work.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Vanden Abeele, V., Wouters, J., Ghesquière, P., Goeleven, A. and Geurts, L. (2015) Game-Based Assessment of Psycho-Acoustic Thresholds: Not All Games Are Equal! *Proceedings of the* 2015 *Annual Symposium on Computer-Human Interaction in Play*, London, October 2015, 331-341. https://doi.org/10.1145/2793107.2793132

[2] Gielis, K., Brito, F., Tournoy, J. and Vanden Abeele, V. (2017) October. Can Card Games Be Used to Assess Mild Cognitive Impairment? A Study of Klondike Solitaire and Cognitive Functions. *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play*, Amsterdam, October 2017, 269-276. https://doi.org/10.1145/3130859.3131328

[3] Adams, W.R. (2017) High-Accuracy Detection of Early Parkinson's Disease Using Multiple Characteristics of Finger Movement While Typing. *PLOS ONE*, **12**, p.e0188226. https://doi.org/10.1371/journal.pone.0188226

[4] Yin, R.K. (1994) Discovering the Future of the Case Study. Method in Evaluation Research. *Evaluation Practice*, **15**, 283-290. https://doi.org/10.1177/109821409401500309

[5] Deng, K., Li, Y., Zhang, H., Wang, J., Albin, R.L. and Guan, Y. (2022) Heterogeneous Digital Biomarker Integration Out-Performs Patient Self-Reports in Predicting Parkinson's Disease. *Communications Biology*, **5**, 1-10. https://doi.org/10.1038/s42003-022-03002-x

[6] Akram, N., Li, H., Ben-Joseph, A., Budu, C., Gallagher, D.A., Bestwick, J.P., Schrag, A., Noyce, A.J. and Simonet, C. (2022) Developing and Assessing a New Web-Based Tapping Test for Measuring Distal Movement in Parkinson's Disease: A Distal Finger Tapping Test. *Scientific Reports*, **12**, 1-11. https://doi.org/10.1038/s41598-021-03563-7

[7] Lopez-de-Ipina, K., Solé-Casals, J., Sánchez-Méndez, J.I., Romero-Garcia, R., Fernandez, E., Requejo, C., Poologaindran, A., Faúndez-Zanuy, M., Martí-Massó, J.F., Bergareche, A. and Suckling, J. (2021) Analysis of Fine Motor Skills in Essential Tremor: Combining Neuroimaging and Handwriting Biomarkers for Early Management. *Frontiers in Human Neuroscience*, **15**, 235. https://doi.org/10.3389/fnhum.2021.648573

[8] Kang, S.J., Choi, J.H., Kim, Y.J., Ma, H.I. and Lee, U. (2015) Development of an Acquisition and Visualization of Forearm Tremors and Pronation/Supination Motor Activities in a Smartphone-Based Environment for an Early Diagnosis of Parkinson's Disease. *Advanced Science and Technology Letters*, **116**, 209-212. https://doi.org/10.14257/astl.2015.116.42

Journal of Intelligent Learning Systems and Applications, 2022, 14, 96-106
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

Intelligent Gaming Input Device for Tilt Recognition

Ilman Shazhaev[1], Dmitry Mikhaylov[1,2], Abdulla Shafeeg[1], Ekaterina Mulyarchik[3]

[1]Farcana, Dubai, UAE
[2]BRICS Institute of Future Networks, Shenzhen, China
[3]School of Psychological Training of Sportsmen "Raise a Champion", Belarussia, Minsk
Email: abdulla.shafeeg@farcana.com

How to cite this paper: Shazhaev, I., Mikhaylov, D., Shafeeg, A. and Mulyarchik, E. (2022) Intelligent Gaming Input Device for Tilt Recognition. *Journal of Intelligent Learning Systems and Applications*, **14**, 96-106.
https://doi.org/10.4236/jilsa.2022.144008

Received: September 11, 2022
Accepted: November 5, 2022
Published: November 8, 2022

Abstract

Farcana has developed a smart a gaming input device, that, apart from being a tool for the gamer to use in the process of gameplay, is also a suitable tool to collect biomedical information about the gamer, which after analysis by the artificial intelligence (AI) system allows informing the gamer about whether the individual is in a state of tilt. Tilt itself is a poor emotional state of the individual that appears due to the latter's inability to control one's emotions in the process of gameplay. The gamer can be either winning or losing, yet the fact that he/she can neither control nor even acknowledge the emotional state is tilt. The latter is an immense factor of impact on the overall success of the individual in the sphere of gaming and one's rating in cybersport. This paper has analyzed numerous studies and patents on the topic at hand. The available literature has provided the necessary insight on the topic of tilt and why it is important to help the gamer acknowledge one's state, especially given deteriorating results. Also, we have proposed a framework for the AI system for tilt recognition.

Keywords

Tilt, Psychological State, Gaming, Cybersport, Progress

1. Introduction

In recent years, esports have been deeply embedded in today's digital subculture among young people. The term "esports" refers to "a kind of sport in which the basic components of the sport are aided by electronic systems", with "the input of players and teams, as well as the output of the system", mediated via human-computer interfaces. Being a professional esports player requires a distinct set of talents and interests than casual gaming.

Esports are competitive to a certain extent. Players need to master not just their physical abilities, but also their ability to work together strategically and keep their emotions in check. Each game has an effect on a player's rank, which is a numerical score or tier that symbolizes a player's skill level. Gamers of all skill levels, from professionals to weekend warriors, take esports very seriously; some take it more seriously than other areas of their lives.

The competitive character of esports games, like that of conventional games and sports, is intended to pique the interest of players by fueling their intrinsic drive to succeed. Sometimes, the level of competitiveness is too high to be considered a fun pastime. Intense emotions are stirred up and social and emotional obstacles are faced as a result of vigorous play. Professional esports need a high level of strategic thinking, the ability to train and improve performance (both individually and as a team), and the ability to keep one's emotions under check.

The success of each game depends on many factors, and psychological state is one of them. Players, who are focused and are confident in their game plan try to stick to it and statistically are unlikely to make a gross mistake that could lead to ultimate failure [1]. Then, it will be much more difficult for the players, over whom emotions have prevailed, to tune in to the result. The condition when a gamer stops controlling himself due to constant failures is called tilt [2]. This paper delves into the issue of tilt from the perspective of psychology and provides a custom-designed gamer solution to the problem in the form of patented technology of a smart gaming input device.

1.1. Research Questions

1) What is tilt and how does it influence the individual in the process of gaming?

2) What methods can be applied to monitor the emotional state of the individual?

3) What factors must one focus on in the process of developing a device capable of monitoring the gamer's emotional state?

1.2. Method

To get a grasp of the topic, one is required to study the scope of literature that was available on the topic. Utilizing the information on Google Scholar and using the Keywords tilt in psychology and tilt in cybersport, one has received 115,000 and 17 responses consequentially all including patents. These have then been categorized and sifted out. The first one chose the time range for the past 5 years to have all studies up-to-date, which left us with 18,200 studies on tilt in psychology and 12 on tilt in cybersport (including patents). After focusing only on full-scope studies made available, we have noted 430 studies on tilt in psychology and 9 on tilt in cybersport (including patents). Additional factors for literature review have been direct relevance to the topic, which decreased the number of studies to 12 studies on tilt in psychology and 6 studies on tilt in cybersport (including patents). These have been assessed with their consequential analysis of the topic at hand. Next, a design of a smart gaming input device is suggested that will help the gamer acknowledge one's state and therefore not harm the latter's gaming.

1.3. Summary

This paper has analyzed numerous studies and patents on the topic at hand. The available literature has provided the necessary insight on the topic of tilt and why it is important to help the gamer acknowledge one's state, especially given deteriorating results.

2. Literature Review

Among all the components that make up the complex structure of the professional competence of an e-sportsman, those professionally significant qualities (competencies) that are directly related to the functional parameters of the ego organism are of considerable importance, *i.e.*, have a psychophysiological nature [3].

In their materials, the authors point out that the professional activity of a cyber-athlete includes long-term training and sports competitions, in the process of which a certain restructuring of brain activity occurs in the cyber-gamer, and the psycho-functional state changes [4]. Neuromarketers recorded objective indicators of changes in the psychofunctional state of gamers with different levels of professional skill. At the same time, the electrical activity of the brain according to the EEG data and the movement of the eyeballs based on the application of the "eye tracking" technology were [5].

An analysis of the emotional state of e-athletes with different levels of professional skill, the intensity of intellectual stress occurring during the game, and the degree of player involvement in the process was carried out based on EEG data [6]. Based on objective consideration of psychophysiological parameters, it became possible to determine those significant qualities of a professional cyber-gamer, which are a necessary resource for the implementation of game tasks during a sports match and can be considered as professionally significant qualities of a cyber-athlete. Such qualities include the ability to allocate attention between gameplay elements that are important for the tactical management of the game [7]. For professional participation in e-sports, it is important to form stable patterns of distribution of attention when studying the screen, which provides the opportunity to pay attention to a greater number of game elements in less time. The ability to control and manage one's emotions, including the skill of self-regulation and management of pre-start stress, and the ability to achieve and maintain the necessary psycho-emotional state characteristic of eustress during competitive periods of a professional career [2].

Studies of the professionally significant qualities and characteristics of e-athletes have shown that at different stages of professional development, during the competitive period, as well as in the course of everyday professional activities, e-athletes face stiff competition from rival teams or individual players, and also experience significant psychological stress and are exposed to various stress factors that prevent the realization of their potential opportunities [4] [8]. A comparative analysis of the psychological state of e-sportsmen and representatives of traditional sports during competitions and at the stage of preparation for them shows that psycho-emotional stress experienced by athletes (fear of failure, experiencing tension problems as a result of a misunderstanding, especially in team sports, etc.) similar characteristics and common psychological mechanisms of occurrence [9]. And in the absence of regulation skills, psycho-emotional stress harms the performance and professional efficiency of athletes in both groups. This condition is called tilt.

2.1. What Is Tilt and Why Does It Occur?

Tilting is a mental condition brought on by unfavorable events and/or emotions that negatively impacts in-game performance and decision-making. Being in a condition of tilt may lead to further tilting, since the negative repercussions induce additional unpleasant feelings. This concept encompasses the internal responses to experiences, the visible deterioration of gaming and decision-making, external and internal influences, and the fact that it is a state of mind.

Tilt is a concept that came to esports from poker. The very meaning is the emotional state of a person, which arises due to constant gaming failures and is expressed by actions that are uncharacteristic of the person performing them [2]. It is easy enough to notice that the player is in a state of tilt. Most likely, the latter will get angry, swear obscene words, pound fists on the keyboard and start provoking the opponents. Such a gamer can no longer control his emotions and actions should he/she choose to continue the game [10]. The game itself in this case becomes an irritating factor, as the gamer cannot succeed or enter the flux in the process of gaming [11]. Therefore, it is better to finish the game, because the probability of further success is extremely low.

In general, tilt is a normal phenomenon, and not a rare occurrence as one may think. No player can constantly demonstrate the same results while remaining unbeaten. Someone loses more often, someone less often, but defeats in games such as CS:GO or Dota are suffered even by experienced gamers. The player's character and one's psychological state are important here (Start, 2019). If one demonstrates a violent reaction to failure, then tilt is just around the corner, with the game ultimately losing its meaning. However, if the player is capable of learning from defeat and maintaining focus, then failures can be replaced by victories [12]. At the same time, players in the state of tilt become easy prey to provocation and various hoaxes on the part of their competition.

Tilt takes place as a result of various factors associated with failures during the game and the inability of the gamer to adequately perceive and respond to them (Garcia-Lanzo *et al.*, 2020). Consequential losses in a depressed psychological state are the result of tilt. If one continues the game after the appearance of tilt, the gamer will make wrong decisions, commit reckless actions and blame teammates for his/her own mistakes (Garcia-Lanzo *et al.*, 2020). The result of such a game will be lost time and a bad mood. Thus, if a gamer feels that one has been overcome with tilt, it is best to just quit the game.

2.2. Types of Tilt

Tilt is expressed in several ways. Conventionally, the tilt state of the individual can be divided into two types, hidden and evident [13]. They depend on what caused the outburst of emotions and what the gamer's character is. Moreover, tilt can be caused not only by constant defeats but also by victories. A gamer who constantly loses gradually loses confidence in his/her abilities and starts to make uncharacteristic mistakes. Yet, with constant winnings, the opposite can occur [13]. The gamer gets an emotional lift and confidence in his/her victory, after which he/she acts more boldly. Yet, in practice, losses provoke tilt much more often than victories.

At first glance, it is difficult to notice the hidden tilt. It is characterized by a depressed inner state of the player. Hidden tilt accumulates due to unexpected losses and can stretch out for several days, weeks, and even months [13]. A gamer who is in a state of hidden tilt will play distractedly, in a completely uncharacteristic manner. Tilt will cause him/her to be confused and completely misunderstand the whole situation [14]. The player will notice the deterioration of the results but is unlikely that one will understand the reason for this, thereby starting to feel depressed. In this situation, it is best to forget about the game for a while and then return to it with renewed strength [14].

In the case of evident tilt, the gamer begins to openly make mistakes and suffers defeats more often. Short-tempered people manifest this via a stream of negative emotions directed at rivals or teammates [15]. If the gamer notices that one is in a state of tilt, one should close the game to avoid both defeat and loss of time.

2.3. Manifestation of Tilt

Tilt can occur at the most unexpected moment of the game. Gamers who, it would seem to be participating in their field of expertise, suddenly begin to lose. At this moment there is a surge of negative emotions [15]. The situation that was just under control is starting to get out of hand. The gamer loses focus and begins to suffer one defeat after another.

At this moment, emotional players begin to lose their nerve. They can bang their fists on the table or start to curse or swear. Calm gamers will show a more restrained reaction to tilt. As a rule, they try to keep their feelings to themselves, but their facial expressions still give away their state [4].

Also, characteristic signs of tilt are:

- the desire to splash out emotions on objects that fall within reach;

- loud insults at opponents and teammates;
- slaps and pounds with one's fists and hands, or even feet on the walls, table, and objects that are nearby;
- insulting opponents in-game chats;
- constant demands will pay off, which will not lead to anything;
- increased aggression during the game, which ends in stupid mistakes;
- dialogue with oneself;
- disgruntled grumbling under his breath.

All these factors negatively affect the gamer's focus and lead to painful injuries.

2.4. The Psychological State of the Gamer

Any emotions affect every completed in-game action, which is why a strong player must keep his cool in order not to make mistakes that affect the outcome of the match [4]. When a person loses, enthusiasm is quickly replaced by anxious thoughts that you do not understand what you should do against this opponent, and your tactics do not work. You start to get nervous and it's even worse to play. But e-sports disciplines are too varied. At the same time, the dynamics of the game process also differ, which strongly affects the emotional state of the player [4]. With limited time to make decisions, a weak player can get lost in his actions and panic, which will negatively affect the outcome of the entire match.

Attitude is also very important. A strong player advises himself not to succumb to in-game provocation, the so-called "BM" (bad manners) [4]. Therefore, every e-sportsman has his variable action plan (game plan) for the game, depending on the given situation [8]. When a player starts to make mistakes and lose, he can also make unreasonable actions that lead to defeat. When you encounter unknown things and don't know the answer to them, you start to get nervous, and it shows. This is a state of body stress [8].

If the causes of rapid heartbeat lie in the pathology of the entire system, then tachycardia is considered serious. The increased work of the department, in this case, is caused by the increased load on this area. It is important to diagnose the disease in time so that the human condition does not become fatal [8]:

- Often the heart rate is observed accompanied by hypertension, as the pressure on the walls of blood vessels becomes excessive, the body will necessarily react to this process [2]. If the arteries are severely affected by atherosclerotic plaques, their flexibility is impaired, and a frequent pulse can lead to the development of a stroke or heart attack [2].
- The pulse represents the oscillation of blood in the vascular bed, which is caused by the process of contraction of the atria and ventricles. When the work of the organ is normal, the blood is ejected from the area of the ventricles and atria, passing into the arteries [2]. If the heartbeat is too active, then the whole rhythm of such blood pumping is disturbed, it stagnates in

the organ, which increases the risk of thrombosis. Bradycardia can also occur against the background of this pathological process [2].

When the gamer is in this state, the latter's self-esteem and moral condition do not allow him to show his/her true level of play: mistakes occur that a person would never make in a normal game.

2.5. Technological Means of Control

As has been made clear when the individual is in a state of stress, abruptly flowing into tilt, the latter's heart rate becomes elevated [2]. There are multiple gadgets from the world of consumer electronics that help people benefit from non-intrusive monitoring. However, these gadgets have never been fixed solely in the gaming industry [2]. Naturally, their primary focus is innovative healthcare [1]. The world has got to know about the Photo-Plethysmography (PPG) technique that is used to analyze peripheral blood volume changes via a near-IR light and a photodetector. The PPG allows receiving information about Heart Rate Variability (HRV) via Inter Beat Intervals (IBIs). The HRV can also help in understanding the user's emotional state [2]. The PPG sensing technology is now present in the form of wearable gadgets, contact-free gadgets, and ordinary devices. The most widespread are wearable sensors most often presented in the form of fitness trackers [1]. The latter are unobtrusive, although are most often used for measuring the heart rate and calories burnt [1].

3. Findings and Analysis

Having multiple smart sensors on the market, one believes that they are not used to their fullest. The potential for implementing the PPG technology in monitoring systems is much higher. Farcana has considered using several autonomic markers responsible for biological measurements of the individual's body. Naturally, from such sensors, one can receive info on the galvanic skin response (GSR), electroencephalogram (EEG), electromyogram, respiration, and electrocardiogram (ECG). Receiving this scope of information would allow us to tell much more about the state of the individual. Of course, these can also measure the skin temperature, and various multimodal data that the developer or gamer himself would like to be demonstrated on the screen for the individual to monitor his/her state. However, to have all these classifiers fulfilled, they must be compatible with the physical aspect of the devices made available. In terms of tilt, Farcana has focused on the heart rate monitoring that allows making the gamer aware of the situation with a further recommendation of action to avoid stress, and consequentially tilt (**Figure 1**). Also, among high-risk cardiac players those who can foresee upcoming tilt can have a chance of preventing cardiac risk/issues/accidents etc.

The available patents have not focused on utilizing the available technology and implementing them in the smart gaming input devices. Thus, the suggested Farcana technology is in no way an infringement of any patents or ownership

Figure 1. Farcana smart gaming input device.

rights of any individual or company (**Figure 1**). We present a device that provides more immersive gameplay combining data from sensors processed by a built-in microprocessor capable of obtaining the gamer's heart rate. This smart gaming input device is executed in the form of a computer pointing device built as a computer mouse with PPG and motion sensors (**Figure 1**). These sensors are attached to the device frame to have an accurate read of data and signals received from the players. In addition, all data are transferred to the cloud for processing and analysis. At the same time, the device system also includes an artificial intelligence system capable of analyzing all data received in real-time. This heart rate data is also used to identify problematic digital biomarkers for various diseases as HRV helps to pinpoint various humoral, neural, and neurovisceral processes to identify various disorders of brain, body, and behavior.

3.1. The Framework

The framework (Wenlu, 2020) for the AI system has main two components: one component is the cloud and the other is the node. The cloud is responsible for storing data as new data is collected, and also to store the trained machine learning models for tilt recognition. Group-based models have been to show greater performance compared to user-specific models and general models, hence in this research we at Farcana Labs train a group-based machine learning model. This model will be updated in the cloud as new data is acquired. The second component of the framework is the node which consists of tilt recognition model, a data processing module and a supervisor for supervising the machine learning process. We also take feedback from the player to confirm the tilt recognition. The two main functionalities of the node are extracting features from the collected physiological signals and using the model trained on the cloud to recognize tilt. **Figure 2** shows a schematic of the proposed framework.

3.2. Training of Group-Based Model

3.2.1. Clustering

If several users are sufficiently similar, a model built for one may be applied to all of them. For this reason, we compare how well each user's model does with the rest of the population and then classify them accordingly. Here's how it works: After creating user-specific models for each subject with the linear SVM algorithm, resulting in a total of n models, each subject's model is evaluated on

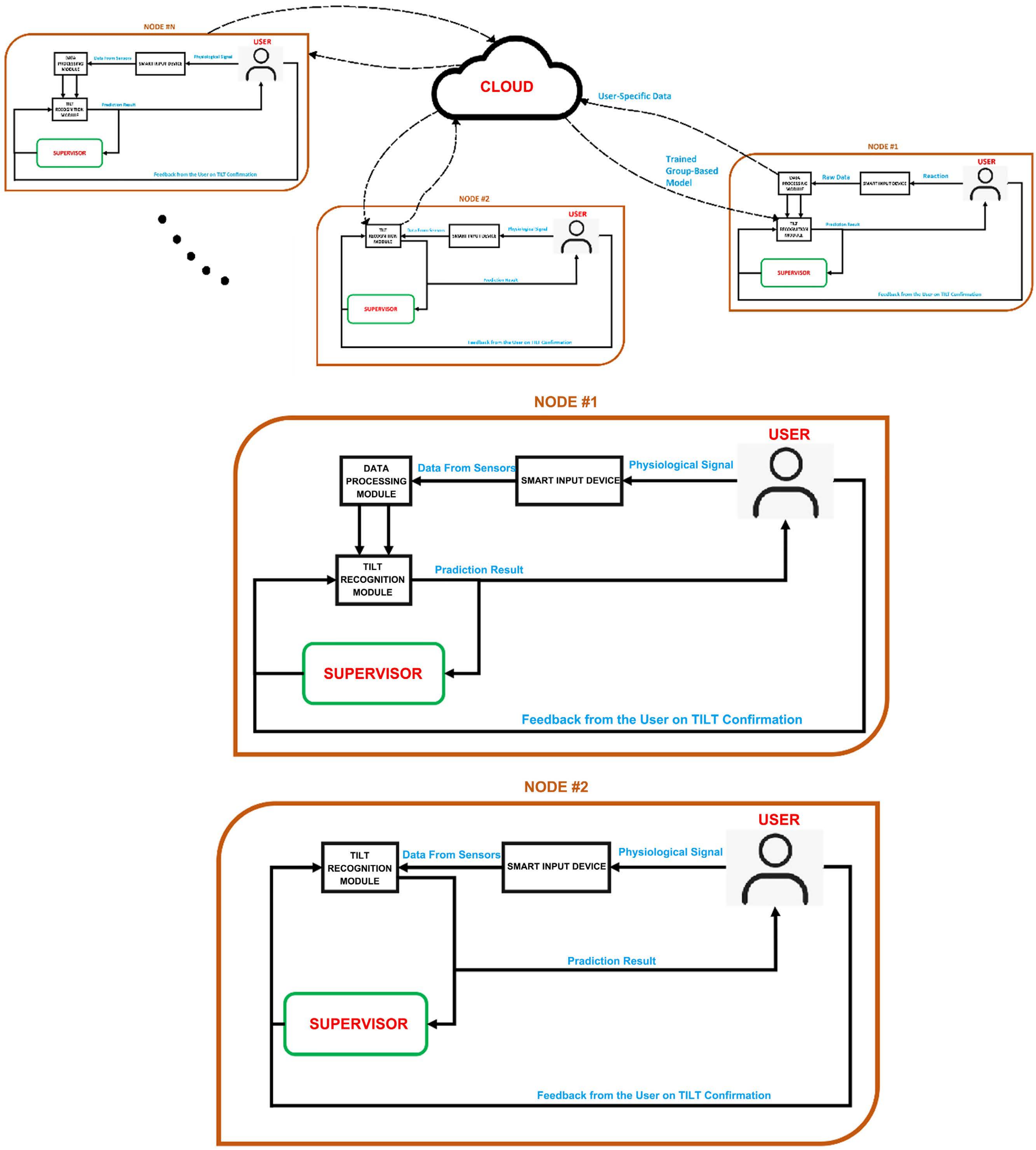


Figure 2. Proposed framework for the AI system.

the other subjects with the F1-score metric, resulting in a n * n matrix, and hierarchical clustering is then used to classify the subjects into groups based on their F1-scores. The model performance matrix P for N subjects may be partitioned into K groups, with each partition holding a list of the subjects that falls into that partition. As a metric of clustering quality, inertia is defined as the sum of sample distances to their nearest cluster centroid.

3.2.2. Training Process

After applying the K clusters, the original data set is partitioned into K subsets, each of which will include all samples from the same topic. Each dataset is used for training the group-based model. The model's predictive capacity is its performance during training. It also indicates the maximum efficiency achievable by the collective model.

3.2.3. Assigning New User to a Relevant Group

For each new user, the optimal group-based model will be the one that can best learn from that user's training data.

3.2.4. Prediction for the New User

A new user's sample data can be predicted after the appropriate group-based model has been identified.

4. Conclusion

Tilt is a state of a player caused by strong emotions from winning or losing, in which the latter plays in a way that is not characteristic of the individual himself starting to make many mistakes in the game. Almost all gamers play to improve their rating, and everyone wants to climb as high as possible, with some succeeding, and others not. The main factor influencing an individual's success is tilt. It prevents the player from thinking about how to play better and because of this, the player stops growing. In many cases, the individual does not even acknowledge this and starts to behave irrationally. Utilizing the available technology of various medical sensors capable of gathering data on the state of the individual, Farcana has developed a smart gaming input device, which not only gathers raw data from the gamer but analyzes it and makes recommendations towards further actions. If the gamer is in a state of tilt, the latter is presented with this information, as well as proof in the form of the heart rate monitor, suggesting that the individual stops. This device is already patented by Farcana and provides the next step in immersive gameplay. We are currently building the proposed AI system, and in future study full details of the system with results of research conducted using this system will be presented.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Lee, J.S. (2021) Exploring Stress in Esports Gaming: Physiological and Data-Driven Approach on Tilt. Doctoral Dissertation, University of California, Irvine.

[2] Smerdov, A., Kiskun, A., Shaniiazov, R., Somov, A. and Burnaev, E. (2019) Understanding Cyber Athletes Behaviour through a Smart Chair: Cs: Go and Monolith Team Scenario. 2019 *IEEE* 5*th World Forum on Internet of Things* (*WF-IoT*), Limerick, 15-18 April 2019, 973-978. https://doi.org/10.1109/WF-IoT.2019.8767295

[3] Fosha, D., Thoma, N. and Yeung, D. (2019) Transforming Emotional Suffering into Flourishing: Metatherapeutic Processing of Positive Affect as a Trans-Theoretical Vehicle for Change. *Counselling Psychology Quarterly*, **32**, 563-593. https://doi.org/10.1080/09515070.2019.1642852

[4] Irwin, S.V. and Naweed, A. (2020) BM'ing, Throwing, Bug Exploiting, and Other Forms of (un)Sportsmanlike Behavior in CS: GO Esports. *Games and Culture*, **15**, 411-433. https://doi.org/10.1177/1555412018804952

[5] Williams, K. (2020) Engaging Games: Studying e-Sports as a Social Emotional Learning Tool. *IEEE Women in Engineering Magazine*, **14**, 27-29. https://doi.org/10.1109/MWIE.2020.3020449

[6] Jeffery, M.J. and Komorous-King R.S. (2018) Method and System for Using Sensors of a Control Device for Control of a Game.

[7] Shulze, J., Marquez, M. and Ruvalcaba, O. (2021) The Biopsychosocial Factors That Impact eSports Players' Well-Being: A Systematic Review. *Journal of Global Sport Management*, 1-25. https://doi.org/10.1080/24704067.2021.1991828

[8] Kou, Y. and Gui, X. (2020) Emotion Regulation in eSports Gaming: A Qualitative Study of League of Legends. *Proceedings of the ACM on Human-Computer Interaction*, **4**, 1-25. https://doi.org/10.1145/3415229

[9] Witkower, Z. and Tracy, J.L. (2019) Bodily Communication of Emotion: Evidence for Extrafacial Behavioral Expressions and Available Coding Systems. *Emotion Review*, **11**, 184-193. https://doi.org/10.1177/1754073917749880

[10] Singh, D. and Sunny, M.M. (2017) Emotion Induced Blindness Is More Sensitive to Changes in Arousal as Compared to Valence of the Emotional Distractor. *Frontiers in Psychology*, **8**, Article No. 1381. https://doi.org/10.3389/fpsyg.2017.01381

[11] Roitblat, Y., Cohensedgh, S., Frig-Levinson, E., Suman, E. and Shterenshis, M. (2021) Emotional Expressions with Minimal Facial Muscle Actions. Report 1: Cues and Targets. *Current Psychology*, **40**, 2133-2141. https://doi.org/10.1007/s12144-019-0151-5

[12] Revord, J., Sweeny, K. and Lyubomirsky, S. (2021) Categorizing the Function of Positive Emotions. *Current Opinion in Behavioral Sciences*, **39**, 93-97. https://doi.org/10.1016/j.cobeha.2021.03.001

[13] Abdullah, S.M.S.A., Ameen, S.Y.A., Sadeeq, M.A. and Zeebaree, S. (2021) Multimodal Emotion Recognition Using Deep Learning. *Journal of Applied Science and Technology Trends*, **2**, 52-58. https://doi.org/10.38094/jastt20291

[14] Meuleman, B. and Rudrauf, D. (2018) Induction and Profiling of Strong Multi-Componential Emotions in Virtual Reality. *IEEE Transactions on Affective Computing*, **12**, 189-202. https://doi.org/10.1109/TAFFC.2018.2864730

[15] Cosentino, S., Randria, E.I., Lin, J.Y., Pellegrini, T., Sessa, S. and Takanishi, A. (2018) Group Emotion Recognition Strategies for Entertainment Robots. 2018 *IEEE/RSJ International Conference on Intelligent Robots and Systems* (*IROS*), Madrid, 1-5 October 2018, 813-818. https://doi.org/10.1109/IROS.2018.8593503

Journal of Intelligent Learning Systems and Applications, 2022, 14, 107-114
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

Video Highlight of Farcana Streaming from Twitch

Ilman Shazhaev[1], Dimitry Mihaylov[2,3], Abdulla Shafeeg[3]

[1]Management Department, Farcana, United Arab Emirates
[2]China Branch of BRICS Institute of Future Networks, Shenzhen, China
[3]Science Department, Farcana, United Arab Emirates
Email: abdulla.shafeeg@farcana.com

How to cite this paper: Shazhaev, I., Mihaylov, D. and Shafeeg, A. (2022) Video Highlight of Farcana Streaming from Twitch. *Journal of Intelligent Learning Systems and Applications*, **14**, 107-114.
https://doi.org/10.4236/jilsa.2022.144009

Received: September 11, 2022
Accepted: November 5, 2022
Published: November 8, 2022

Abstract

The Highlights are the most interesting, selling moments from video stream, which can make the viewer watch the entire video. They are like a shop window: everything that is bright and colorful goes there. Seeing them, the user can understand in advance what is inside the video. And they are more versatile than trailers: they can be made shorter or longer, embedded in different places in the user interface. The user sees a selection of highlights as soon as he gets to the website or watches a video clip on YouTube or even a section of a stream of a popular blogger/influencer. The user's attention is immediately attracted by the most memorable shots. Naturally, it is becoming more tedious to manually create all video highlights, due to the immense amount of material that the highlights are needed for. Thus, creating an algorithm, capable of automating the process would make the process significantly simpler. Besides easing up the work, this process will pave the way to a whole set of new applications that before did not seem real. At the same time, this would be a new process where the AI would not need to be fully supervised by a human, but be capable of identifying and labeling the most interesting and attractive moments on screen. After doing a literature review on video highlight detection, this paper has utilized the model presented in the study by [1] to determine the possibility of attaining highlights from the Farcana 2.0 version Twitch video feed.

Keywords

Highlights, Intelligent System, AI, Algorithms

1. Introduction

One cannot say that creating an automated AI algorithm to decrease the level of supervision during the process of highlights creation has not been attempted before. On the contrary, numerous scholars have attempted and succeeded to make the highlights gathering process much easier. This paper focuses on the study attempted by [1]. Scholars have made it their priority to develop a way to automate the process of having the AI automatically gather the most interesting moments in on-screen production. Their method of contrastive learning is undeniably complex, yet it has proven to be a success. Other scholars have attempted studies mostly focusing on solutions that required full-time supervision of the process by a human being. Yet, in study [1], they have managed to achieve an automated solution decreasing the amount of supervision necessary. Their framework encoded the video "into a vector representation" [1], thereby teaching the AI to pick the most interesting moments that could be classified as video highlights and further demonstrated to a large audience. Their method was compared to the existing solutions requiring state-of-the-art approaches and equipment heavily relying on substantial input from external data. Yet, the approach in [1] provided us with a simpler solution requiring much fewer resources than say, the LM system which employs as many as 10 million videos in the process of training their algorithm.

This algorithm can be applied in video retrieval, finding recommended content, help in browsing, and naturally editing. Although this particular paper is focused on setting the task of receiving the best highlights of the Farcana streaming on Twitch utilizing the gameplay of the Alpha 2.0 version release. The application of method presented in [1] allows us to update the innovative approach and automate the process of video highlight retrieval. This saves on manpower and time, at the same time saving on costs.

What happens in the frame, whether it's an explosion, a gunshot, or a fatality move, affects the conversion in different ways. This influence can be measured—some statistics allow you to guess in advance which fragments of the film are best suited for the role of highlights [2]. And then special slicing algorithms are turned on part of a large system for preparing content for streaming. It can create a highlight from any video file and at any time after it enters the system. The operator creates a job to generate a highlight, and the system processes it according to the set priorities.

There are many algorithms. We have used several in our experience. The first is the scene detection algorithm. He must determine the beginning and end of the desired piece of the film. In the course of training, the algorithm was first shown various shooting angles, taught to find them, and then, from the selected angles, they were taught to highlight specific scenes. Also, information from the video and audio stream is used to determine suitable segments of the film. This is necessary so that dialogs are not interrupted when creating highlights.

The second is the scene selection algorithm for creating a highlight. We trained our algorithm based on Twitch streams of Farcana gameplay. We took all the streams that were available in the Alpha 2.0 version released in the second quarter of 2022 and available only under Private Sale terms. Now, having a trained algorithm, we can rank any scene of a new stream file, calculating the probability of how this scene is suitable for creating a highlight.

When the set of scenes is ready, the algorithm selects the final candidates, from which the highlight will turn out. At the same stage, one can do image processing, color correction, and filters. The results are reviewed by a post-moderation person. The latter can always reject a ready-made highlight if it turned out badly. If the highlight passes the test, it is placed on special servers, from where it is delivered to users using a CDN.

2. Literature Review

Video Highlight Detection has gone through a substantial path of development. Of course, initial attempts were made in the sphere of sports with a focus on sporting highlights. Then this trend moved forward to social media and first-person feeds [3]. However, at the start of this path most methods applied were focused solely on the heavily supervised approach having an individual moderate the content and general feed.

Further steps in video summarization have started to become predominantly unsupervised with scholars relying on heuristics and representativeness to be able to obtain a summary [3]. At the same time, weakly supervised methods have been developed for the production of video annotations enabling users to benefit from video tagging. Nonetheless, a substantial number of scholars have attempted to shift progress in R&D on unsupervised models of highlights capture. For example, [4] have acknowledged that AI is at the forefront of technological development and the shift to a new level of global operation. These scholars have noted the immense success of WSC Sport as a sports algorithm framework capable of providing highlights of NBA team statistics, highlights, and various kinds of data to sports fans in real-time. An automagical solution is what WSC Sport calls themselves and explains with the definition of the word from a Dictionary: something happens automatically, but from the outside, no one can explain exactly how. The initially Israeli startup WSC Sport broke into the American market in late 2015 and turned the NBA's highlighting process around [5]. There, video reviews are especially needed, because dozens of effective actions are performed in each game, and a fan needs to highlight the main ones.

WSC Sports Technology is an artificial intelligence that can collect highlights from really key moments without any human intervention at all. The technology recognizes the signs of the most important episodes: how the score changes after an accurate throw, how emotionally the players and fans react, whether the commentator powerfully intones, and so on [4] [5]. During the match, the algorithm generates a rating from all game segments in general, each of which is marked with details: the characters, the type of throw (long-range, medium, from under the ring, with resistance or open), the features of the episode (quick attack, positional attack, throw immediately after transmission) [4] [5]. The technology allows you to create highlights at any time and for any duration, just set these parameters. For example, cuts of the actions of a particular player fit perfectly into a live broadcast, if he is especially (or unexpectedly) good, in a long break or immediately after the match, a video of the main moments as a whole is organic. A club or an owner of media rights can use customized highlights at any time and for any reason: if one wants to highlight the main antagonist of a winning match—the algorithm collects a minute video for Instagram [5]. If the focus is to spectacularly announce the extension of a contract—in a couple of clicks one can receive the best moments of the player's career from WSC Sport.

Naturally, it was possible to do the same thing before, but manually and with a constant risk of missing something. Israel's technology, which is rapidly learning itself and already conquering other sports, is a perfect example of intensive development when a standardized task is transferred to smart development, and human resources can be directed to creativity [5]. In addition, manual assembly takes extra time, and WSC Sport returns the advantage to clubs and the league in the fight for the second screen of the viewer.

There is no exact data or research details on how this algorithm works, but the scientists say their technology is constantly learning, so the further they go, the better they can harness the full power of artificial intelligence [5]. The algorithm prioritizes in real time because different events can be the most important in different matches. The most important thing is not just to collect moments in a cut, but with the help of AI technologies to succinctly formulate the same plot that has developed in the match [2].

WSC Sport is one of the flagships in the creation of ultra-personalized content. In addition to the NBA, the American football league MLS, the most progressive sports media in the world, [2], and others are already cooperating with the company. They all understand that in parallel with global coverage, the demand for individual media consumption is growing. Among the audience, many want to determine for themselves what exactly and when exactly to watch [6]. Instead of targeting content to specific markets or demographics, one can now give users the chance to customize it for themselves [2]. Integration with a chatbot in a messenger is unlimited access for a regular user to a huge content database, which is automatically (or rather, automatically) created for any occasion. Now any media company and organization that wants to be active and visible through their content has a powerful additional tool.

Yet, the WSC Sports technology seems to be too complex and is based on a substantial data-gathering process. A new solution has been suggested by [1]. Scholars have utilized the contrastive learning methodology. The main focus of their framework is to teach the algorithm an unsupervised scope of actions that can be applied for multiple downstream tasks. The AI is trained to create a random mapping, localization, identification, and segmentation of images in a video stream. The algorithm also performs the cropping and editing functions presenting the best images in a given sequence thereby transforming the original frame to avoid spoilers and increase attention and awareness of the specific product. The scholars have initially considered their methodology with a focus on contrastive learning to conduct the visual transformation in the most straightforward way possible. By applying the dropout sequence as part of the contrastive learning procedure, [1] have managed to develop an unsupervised sequence embedded in the algorithm that is now capable of a general transformation with the highlight presentation as the outcome. Scholars have developed an unsupervised framework that is capable of picking the most attractive clips that further make up a single vector. Then the dropout sequence is applied for the transformation to take place and to be able to map the two vectors, localize the moment of interest, and cluster the imaging, with its further segmentation. As presented in **Figure 1**, the framework allows us to cluster the video based on content, yet with the dropout fluctuation and change. This allows for better clusters to be chosen for further highlighting and production of video embedding. **Figure 1** shows the same video content embedded in vectors with the non-highlights (left) cluster that have been dropped, and the highlights (right) cluster that the algorithm has decided to retain for further operation and editing.

Figure 1 clearly shows that the non-highlight clusters do not fit the framework as they overlap, are inconsistent, and are different in both scope and content. Each video feed is presented as a different color, with the application of random dropout sequences to categorize the feed. This demonstrates the difficulty there is in directly utilizing the video feed and using it in the video highlights compilation from the start.

3. Suggested Approach and Results

3.1. Overview

The approach presented in [1] suggests first splitting the video feed into same-length clips. Each clip is represented using a vector, whereas the whole video is presented in the form of a clip sequence. Then the C3D action recognition model is implemented to extract the raw features of dimension. The main goal here is to achieve high scores through a contrasting learning framework (**Figure 2**).

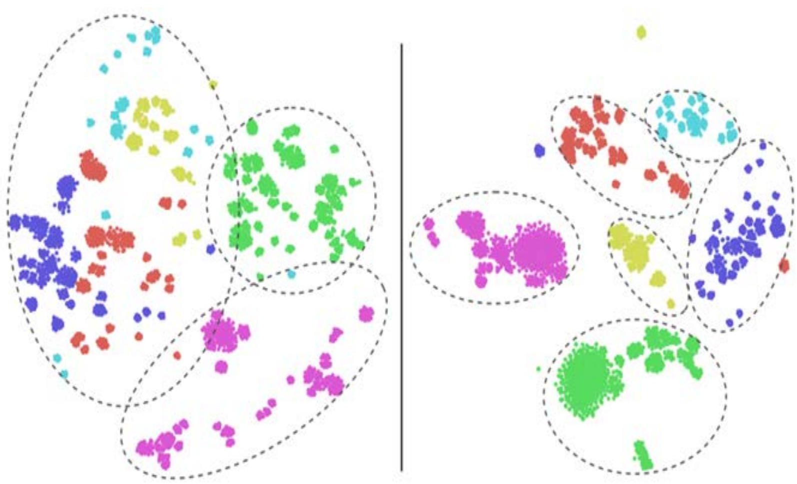


Figure 1. The non-highlights clusters vs highlights clusters [1].

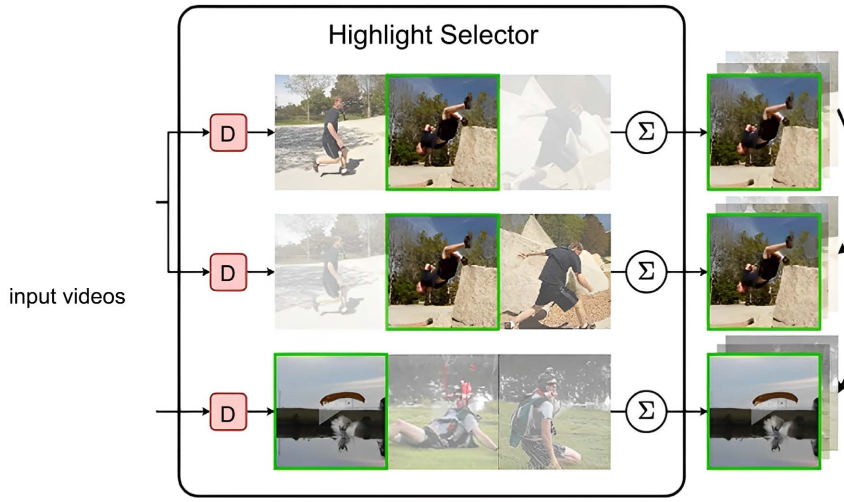


Figure 2. C3D action recognition clip selector.

Here we learn to map the changed vectors farther away from each other in the sequence. The clip selector establishes a specific attention score further used to compute a specific sum of the clips measured against the scale of the highlight clips.

3.2. Generic Attention Layer + Highlight Selector

The generic attention layer is an important factorial inclusion to make the notation as simple as possible. Figure 3 demonstrates the peculiar impact that this layer has on the transformation process of the entire perception of the video feed. The attention layer is defined as such that queries Y using X. At the output, we receive a generic attention score in the form of an attention map with a particular interdependence level between x_i and y_j. This allows us to determine the interaction and the level of interchange between the various sequences presented in the video feed. This is also required for the smooth transition of the audio feed between the transformed vectors. After all, if the query and key highlight sequences are different in length, that would mean that there would be a residual part requiring us to remove the clip from the framework altogether (as in Figure 1).

The Farcana video feed from the Twitch clip was broken into an equal number of clips with a predetermined number of frames for each feed. As a result, we obtained N clips. As has been mentioned each clip is passed through a pre-trained C3D action recognition model, to size each clip. The generic attention layer then models the relationship between clips to determine their coherency and their compatibility. Then the second sequence of generic attention layer identification is conducted to determine the specific features of the query pool. As a result, we receive a set of attention scores with a comparison to the highlights score. These scores remain unsupervised in the process of embedding with the application of contrasting learning. The highlight detection is aimed at capturing the relationship between different clips, even if they are not from the same moment in time of the Farcana Twitch feed.

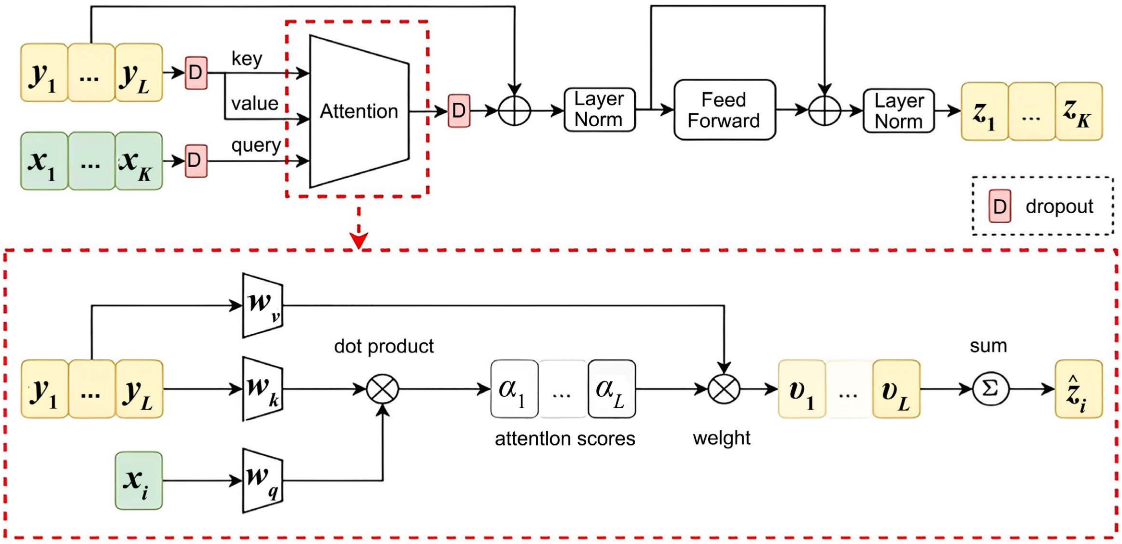


Figure 3. Generic attention layer.

3.3. Analysis

The approach in [1] has allowed to teach the algorithm the process of video clustering bearing in mind the ongoing dropout fluctuations. This model learns to pick out moments that cannot be considered highlights due to the peculiar dropout moment and a low generic attention score. Yet, if the model does foremost drop the non-highlight clips, it may have been safe to assume that the remaining would be safe to use for the highlights. Yet, this is not the case, as should the algorithm stop at this moment, the quality of clusters will deteriorate, and therefore the performance of the contrastive learning model suggested by [1] would simply become irrelevant and inapplicable in any potential sphere of operation.

Instead, one suggests resetting the algorithm from start by comparing the non-highlight clips to the remaining hypothetically highlight clips. We use the C3D Action recognition model on both, having both groups of clips pass the whole process again to receive the output of the first generic action level. Again, the remaining clips are sampled into non-highlights and highlights. This cycle is repeated another twenty times with a different set of dropout requirements. The embedding outputs should form a separate cluster, with a specific intra-cluster distance. The closer the video embeddings are to each other the greater interrelationship between them. Whereas a greater distance would demonstrate that the videos are too far apart from each other both in terms of content and even generic attention scores. To measure the distance, the cosine distance is used. This score allows the algorithm to receive intra- and inter-cluster statistical data and consequentially compute the mean value across all videos.

The highlight clips form a tight cluster with embeddings located close to each other, whereas those at a greater intra-cluster distance identify the clip to be a non-highlight clip. In addition, as is visible from **Figure 1**, the highlight clips form separated clusters that are located at a substantial difference from each other. This allows the algorithm to pinpoint the best moments and further edit them into a highlight clip.

4. Conclusion and Limitations

This paper has utilized the model presented in the study by [1] to determine the possibility of attaining highlights from the Farcana 2.0 version Twitch video feed. We have implemented a simple unsupervised method that is repeatable and does not require additional complex state-of-the-art equipment or method. This model distinguishes the dropouts from other models, with a highlight received as the outcome. No external large data is collected, with the model having the possibility to remain unsupervised/weakly supervised in highlight detection.

Despite the evident success, there are certain limitations to the process. It has been determined that the non-highlight clips do not contain the same amount of information, yet there is still a chance that this can be disproven over a higher number of sequences running the model. A possible explanation could be the substantially different content of the videos. A possible solution is by choosing video feeds that are closest to the specific video in question to form a topical constitution of negative pairs. The current study only chose random feeds of Farcana 2.0 version release gameplay.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Badamdorj, T., Rochan, M., Wang, Y. and Cheng, L. (2022) Contrastive Learning for Unsupervised Video Highlight Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, LA, 18-24 June 2022, 14042-14052. https://doi.org/10.1109/CVPR52688.2022.01365

[2] Hong, F.T., Huang, X., Li, W.H. and Zheng, W.S. (2020) Mini-Net: Multiple Instance Ranking Network for Video Highlight Detection. *European Conference on Computer Vision*, Springer, Cham, 345-360. https://doi.org/10.1007/978-3-030-58601-0_21

[3] Jiao, Y., Li, Z., Huang, S., Yang, X., Liu, B. and Zhang, T. (2018) Three-Dimensional Attention-Based Deep Ranking Model for Video Highlight Detection. *IEEE Transactions on Multimedia*, **20**, 2693-2705. https://doi.org/10.1109/TMM.2018.2815998

[4] Campbell, C., Sands, S., Ferraro, C., Tsao, H.Y.J. and Mavrommatis, A. (2020) From Data to Action: How Marketers Can Leverage AI. *Business Horizons*, **63**, 227-243. https://doi.org/10.1016/j.bushor.2019.12.002

[5] Liu, W., Yan, C.C., Liu, J. and Ma, H. (2017) Deep Learning-Based Basketball Video Analysis for Intelligent Arena Application. *Multimedia Tools and Applications*, **76**, 24983-25001. https://doi.org/10.1007/s11042-017-5002-5

[6] Ye, Q., Shen, X., Gao, Y., Wang, Z., Bi, Q., Li, P. and Yang, G. (2021) Temporal Cue-Guided Video Highlight Detection with Low-Rank Audio-Visual Fusion. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Montreal, 10-17 October 2021, 7950-7959. https://doi.org/10.1109/ICCV48922.2021.00785

Scientific Research Publishing

Journal of Intelligent Learning Systems and Applications, 2022, 14, 115-131
https://www.scirp.org/journal/jilsa
ISSN Online: 2150-8410
ISSN Print: 2150-8402

Gamification in Education: A Study of Design-Based Learning in Operationalizing a Game Studio for Serious Games

James Hutson, Ben Fulcher, Joseph Weber

Lindenwood University, Missouri, USA
Email: jhutson@lindenwood.edu

How to cite this paper: Hutson, J., Fulcher, B. and Weber, J. (2022) Gamification in Education: A Study of Design-Based Learning in Operationalizing a Game Studio for Serious Games. *Journal of Intelligent Learning Systems and Applications*, **14**, 115-131.
https://doi.org/10.4236/jilsa.2022.144010

Received: October 21, 2022
Accepted: November 13, 2022
Published: November 16, 2022

Abstract

The gamification of learning has proven educational benefits, especially in secondary education. Studies confirm the successful engagement of students with improved time on task, motivation and learning outcomes. At the same time, there remains little research on games and learning at the postsecondary level of education where traditional pedagogies remain the norm. Studies that have been conducted remain almost exclusively restricted to science programs, including medicine and engineering. Moreover, postsecondary subject-matter experts who have created their own gamified experiences often are forced to do so on an *ad hoc* basis either on their own, teaching themselves game engines, or with irregular support from experts in the field. But to ensure a well-designed, developed, and high-quality educational experience that leads to desired outcomes for a field, a sustainable infrastructure needs to be developed in institutions that have (or can partner with) others that have an established game design program. Moreover, such a design-based learning approach can be embedded within an existing studio model to help educate participants while producing an educational product. As such, this qualitative case study provides an example of the process of operationalizing a game design studio from pre-production through post-production, drawing from the design and development of the educational video game *The Museum of the Lost VR* (2022). The results, resources, and classification system presented are scalable and provide models for different sized institutions. Methods to develop a sustainable infrastructure are presented to ensure interdisciplinary partnerships across departments and institutions with game design programs to collaborate and create educational experiences that optimize user experience and learning outcomes.

Keywords

Gamification, Game-Based Learning, Pedagogy, Game Studio, Project

Management, Team Structure

1. Introduction

The possibilities provided through the gamification of education are becoming more broadly accepted beyond secondary education. The softening has been achieved through gaming being adopted in many areas of daily life for users from a variety of backgrounds and ages. When applied to education, gaming can ensure learning materials to become more engaging and motivating. Such examples already include Khan Academy and applications like Duolingo. The same methodologies used in these examples are being used more and more in the virtual classroom, as well [1] [2]. One reason that gamification is being adopted is that active and participatory learning strategies are superior to passive alternatives to engage learners. Recent studies have demonstrated that there is greater retention and engagement in the learning process if there is purposeful participation included in lesson plans. Dastyar [3] confirms that the role of motivating factors that work in tandem with participatory learning increases both motivation and academic achievement in students. At the same time, games are also emblematic of a broad cultural shift toward said participatory culture and offer ways to model the same educational experiences. For instance, the act of gaming and gaming communities move players/students beyond passively consuming media/information to actively participating and producing an experience through their interaction with the game itself and other players. Moving beyond playing games, the process of designing, evaluating, and developing said experiences itself can be an intentional educational strategy through design-based learning (DBL). Students involved in the creation of gaming experiences have the ability not only to add to their portfolio but gain valuable skills for a variety of industries in their role in a team for a game studio. The process aligns closely with the ADDIE model for instructional design (**Figure 1**). The same models for DBL as a pedagogical approach and embracing participatory culture may be adopted by educators as in this study [4] [5] [6].

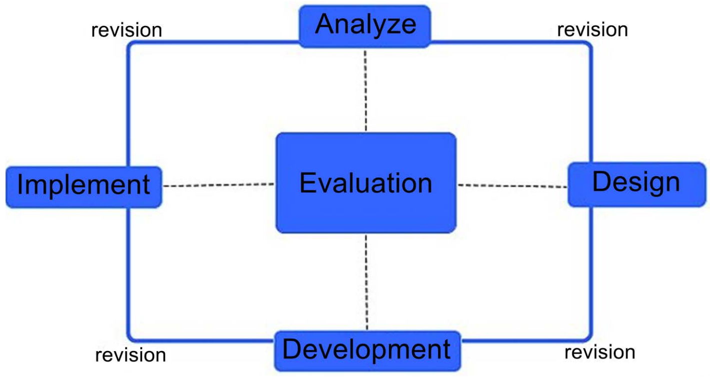


Figure 1. ADDIE model of instructional design.

There are various strategies educators at different levels can take when gamifying their curriculum. For instance, activities within the classroom itself may be gamified, such as adding point systems to responses and discussions to full immersive adventure experiences that last an entire course [7]. Others who wish to leverage the benefits of video games to engage and motivate students, improve their overall experience in the course [8]. Regardless of the resources of an educator's institution, there are many freely available educational video games on the web, such as ABCYa with experiences from pre-K to grade 6. Similarly, serious games may be found on FUNBRAIN and Education.com through libraries of games, videos, and books for K-8 students. Yet, while there are many existing resources available, especially with regards to secondary educational gaming, postsecondary educators are left with few choices aside from the publisher resources included in e-textbooks, such as interactive readings and self-quizzing options. Gamified experiences tailored to various disciplines remain out of reach for most college students who would benefit from a more participatory approach to teaching and learning. Studies on the positive impact of gaming for college students often look at existing video games and consider how different genres (role-playing games [RPG], first-person shooters [FPS], etc.) may support developing certain skills, such as critical thinking or teamwork [9]. The ability to create an intentional learning experience to support a specific outcome designed by a subject matter expert (SME) in the field not only leads to a better understand of the topic being covered, but also can more readily align with the additional durable skills on which current research focuses [10] [11] [12] [13]. But subject matter experts, no matter how accomplished in their fields, cannot accomplish this task alone; an infrastructure, supported by constituents both internally and externally, must be designed and implemented.

One strategy to address the issue of providing support for faculty wishing to create a gamified experience and serious game for their courses, while also giving students design thinking experience, is for institutions to develop a sustainable infrastructure built on the substructure of a game design program [14] [15]. Instead of merely having a SME work with a team on an educational product, the experience itself can be a learning experience and extend beyond designers and coders. Students in the discipline of the SME along with those from game design can collaborate using DBL for project-based learning to gain skills just-in-time (JIT) as the iterative process of design unfolds. With an internally funded grant, a team of faculty, graduate students, contractors, and consultants conducted a study on the potential of such a model. The following treatment presents the results of the study to operationalize an educational game studio at a mid-sized, private Midwestern university through the creation of the video game *The Museum of the Lost VR* [16]. The process of designing a successful production schedule and team structure is presented with recommendations for all phases from pre-production, production, post-production, and, finally, distribution. While the results from the case study are ideal for institutions with a game design department with a graduate program, the design is scalable and provides models for institutions of various sizes and resources. Furthermore, the pedagogical approach of DBL in the process of gamifying educational experiences provides a model for all types of gamification lessons.

2. Literature Review

A broader variety of learners have been engaged as of late using educational games and the gamification of instruction in the classroom [17]. The use of game-based instruction is certainly not a recent phenomenon, but the broad adoption of recreational gaming for a more heterogeneous population along with ready availability of the technologies to create gaming experiences have coalesced to prompt adoption of serious games for educational purposes, as well [18] [19]. The watershed moment has arrived at the ideal moment as educators struggled and failed to engage students in the post-COVID era through many motivational strategies. The standard extrinsic motivation strategies employed in education, unfortunately, are effective only for a short time. On the other hand, gameplay can sustain attention and engagement for longer periods [17].

2.1. Theories for Gamification in Learning and Education

The major theories of learning when considering gamifying education may be broken into four categories. Motivation, Self-determination, Achievement Goal, and Social Learning or Situated Learning theories may all be considered when considering a gaming strategy, however, the one selected will depend upon the field, lesson, activity, or outcome desired. As an example, should motivation be the most significant outcome, educators may consider Motivation Theory for the experience. Studies have borne out that motivation is the most important element for a successful learning experience in a gamified environment [20]. Self-Determination Theory is also relevant as it builds on the dual system as a macrotheory of motivation [21] [22]. Building on the extrinsic motivational theory, self-determination assumes that a learner's motivation is influenced by the immediate learning environment, and both social and cultural factors need be considered. In order to further develop inborn psychological desires to learn and achieve, educators may offer several options during instruction to ensure several variables are addressed. These variables include the perception of students that they have the ability to complete an activity, feel in control, and enjoy a sense of community with other learners.

Turning to Achievement Goal Theory, the related macrotheory of gamification takes for granted that learners are motivated by a desire to complete a task or reach a specific goal. [23] [24]. Given the interest in achievement, the theory has two specific goals which include performance and mastery [25] [26] [27]. The more self-motivated a learner is, the more they will appreciate and relate to mastery goals. These goals relate to an innate desire to understand a concept, finish a task or acquire abilities. Alternatively, learners who look to outperform others at the same task would fair better with performance goals, as these are "ego-involved". These learners are more interested in comparing themselves to others in a social setting than the actual learning outcomes associated with the task [28]. Along the same lines, Social or Situated Learning Theories also consider the significance of social motivations and learning environments. However, instead of competing with others in a social setting, these learners do better observing others and their behaviors [29]. In order to be successful, educators need to first identify a student to act as a model for others to observe and then model their own behavior on once witnessing the outcomes of their actions.

2.2. Game Design Studio

However, before implementing these considerations of student motivation, educators need additional technical and operational support. Gamifying an experience requires both subject-matter expertise and a technical knowledge of game mechanics and other specialized skills if developing video games. An example of an educational institution addressing these concerns, and the operational and workflow issues inherent in developing educational video games is the Massachusetts Institute of Technology (MIT). The approach took the shape of a Comparative Media Studies (CMS) project that aimed to consider real-world challenges through academic theories of games. Given the project's success, MIT further expanded in 2006 with the establishment of the Singapore-MIT GAMBIT (Gamers, Aesthetics, Mechanics, Business, Innovation, and Technology) Game Lab. Through "applied humanism", games are created within an academic context. The goal of the project was to bring academics and researchers "down from the ivory tower" to demonstrate the values of their theories through building games ([30], p. 256). In the summer of 2007, 30 students from Singapore were brought for a nine-week internship to work with MIT graduate and undergraduate students to create six games. Each team had seven members: two programmers, two artists, a game designer, a test lead, and a project manager. In addition, a two-person audio team provided sound and music support for all development teams. Given the constraints of time, top-down oversight was not possible, and so the "Scrum" project management model was adopted from industry. This model requires agility in project management and acts on new findings, unexpected outcomes, and user feedback quickly and efficiently. Teams would iterate and develop a playable build of their games every two weeks and receive feedback. The goal in this instance was for each team to demonstrate a single research idea in their game project [30].

Another example uses a more open pedagogical approach. 038 Games (https://038games.nl/), a serious gaming studio, works with clients that bring a specific problem or learning outcome that needs to be addressed through gamification. The company offers a minor and educational experience open to anyone who wishes to enroll from around the globe. In 20-week cohorts, students with various backgrounds work on as many as 15 different projects annually. The role of the instructor in the experience is as a coach as they do not provide lecture content. Instead, the instructor serves as a project manager for the team. Team building activities are included in the course requirements for the minor. For instance, coders are set to work on the UI elements while art and design students begin work on the concept art for the game. Examples from the studio include an escape room, which can include a variety of discipline-specific information, and players are then encouraged to work together, learning new information in order to exit the room. The collaborative interactions between researchers and subject matter experts (SME) and game designers and developers in the two examples above provided a starting point for our study in considering the logistics of such an enterprise.

3. Methods

3.1. Game Design & Production Process

3.1.1. Team Structure

The roles outlined in the grant proposal ensured faculty and student participation, along with outside contractors to fill any skills gap. The process followed recommendations for game development and documentation [31] [32] [33] [34]. The resident art historian served as subject-matter expert (SME) and primary investigator (PI) for the grant. The SME created the Game Design Document (GDD), preliminary level design, recorded audio clips, provided text-based information to be included on pedestals and for quizzes. The resident game design faculty member acted as project manager and oversaw meetings, kept updated files in a shared cloud drive for the team, and provided trouble shooting for team members attempting new functionality in Unity. Two game design graduate students served as game and level designers for the project; an external coder was contracted for the UI/UX elements, such as the functionality of the main menu and quizzes; and, finally, a consultant who had developed educational historical games advised on process. Only the two faculty members could meet physically, and thus the project team needed to adopt the strategies of virtual (remote) teams (**Figure 2**) in dimension of time, space, and culture [35]. In fact, the entire team never met together physically. The benefits of such a team include flexibility and a broader skillset, though motivation need be considered [36]. At the same time, the management structure was more formal given the requirements of the project and expertise required of the team to utilize the decentralized virtual team model. As the project required both expertise in the game engine, asset creation and management and development as well as content expertise in the material being covered, the matrix team model was adopted for logistical purposes. In the "two-boss system" (**Figure 3**), individuals report to different managers for various aspects of work due to expertise. Throughout the project, the team met with both project managers and received feedback and instruction based upon the expertise needed relating to design, mechanics, debugging, and instructional material. As such, a hybrid team structure was developed that used both (**Figure 4**) the "two-boss system" and the virtual communication strategies for the project.

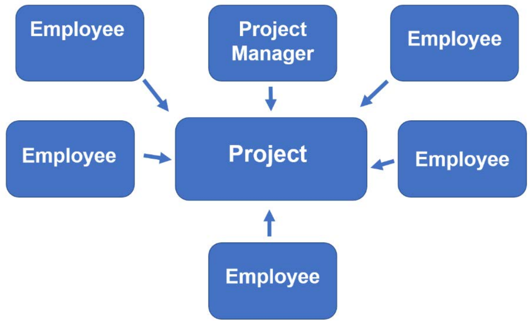


Figure 2. Virtual team diagram.

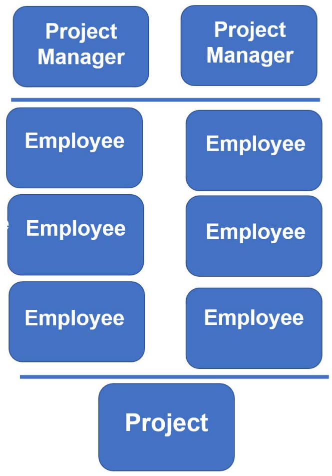


Figure 3. Matrix team diagram.

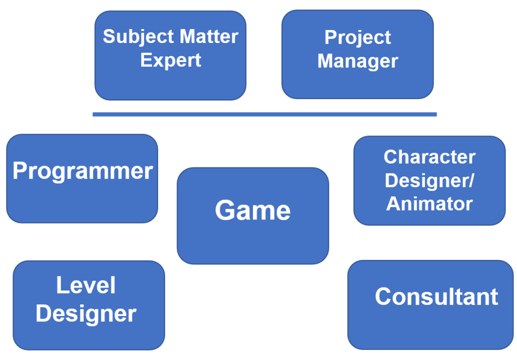


Figure 4. Hybrid team diagram.

3.1.2. Pre-Production

The initial design and production of the educational game was divided into three phases-pre-production, production, and post-production. The project ran from January-April of 2022. After securing internal funding, a pitch deck was developed in early January, which served as the GDD for the project. The document served as a centralized vision for the project and helped communicate expectations to team members, provide context and background for the project, and outline necessary game mechanics and components for the project [37]. During the pre-production process, planning for the project was completed and preliminary documents were constructed, including the GDD. The preproduction phase saw team members determine what tools and versions would be used for the project development. Unity game engine was decided due to team familiarity. During this phase, sprint meeting times were decided as well as communication platforms (Microsoft Teams and Discord), task delegation, and individual expectations. In this phase, team members were selected, and the game's core play loop was defined.

3.1.3. Production

During the production phase, work in-engine began with team members being relegated tasks and game mechanics construction began. During this roughly twelve-week phase, the team worked on modeling or finding game assets, designing game mechanics, developing User Interface (UI) components, and general troubleshooting. This phase also saw the widest delineation from the predetermined GDD sprint schedule. Regular spring meetings were conducted every week-every other week for 15 minutes and a full hour report alternating weeks. During the meetings, each team member reported on their activities for the past week and had questions answered on the educational goals of the games and identified technical solutions for player engagement. All meetings were conducted synchronously via video conference platform, and, in fact, the team never met physically being located in different regions around the United States. In general, the early stages of the project found that the development team was completing tasks much quicker than initially outlined in the GDD. During the first two sprint meetings, all expected tasks were completed before the outlined dates in the GDD. Because of this, the development process saw a shift in workflow towards the end of the expected timeline. Basic locomotion and game mechanics were developed early on and provided more time later in the phase for UI creation and fixing game elements that were not working. By the end of the production phase, the team had developed working mechanics for all project Minimal Viable Product (MVP) goals.

3.1.4. Post-Production

Once all levels had been completed and a beta version was ready for testing, the team oversaw recruitment and distribution of the build for the playtesting. In this final phase of post-production, marketing content was developed in the form of a copy and a trailer for the game, as well as final touches were to be made in the game. A period of playtesting lasted two weeks in which participants would complete a survey after playing the game to facilitate critique and feedback. These surveys would be used to pinpoint bugs and problems in the game for fixing during this period. Game platforms such as Steam and Itch.io were also reviewed for future submission. This phase also saw the game project being accepted to a virtual conference dedicated to gamification in education.

4. Results

The challenges faced during the design and development of the serious educational game reflect that in the literature [38]. To improve upon the development and evaluation processes, the following recommendations and insights were gleaned from the team at the close of the project. For instance, the pre-production phase should have been preceded by a finished GDD and Level Design Document (LDD). The pitch deck was created after the team was assembled, and a more efficient design process would be made possible with a fully fleshed out project prior to the development cycle. For such a visually rich project, reference images and examples should have been readily available from the outset for the whole team. To facilitate this process for a PI and/or SME, a template should be developed for internal use that can be filled out through backwards learning design. Additionally, game development documentation and institutional development policy should be created unique to the organization's needs [39]. Creating an expanded onboarding package for new SME would also be useful in what to consider, much like a design document in working for a client. The training should also include building expectations early on in the timeline so that the SME avoids continually adding additional functionality and UI elements as the project progresses.

In the pre-production phase, recruitment should be of paramount concern and preparations should be made to ensure the most qualified candidates can be identified efficiently and on the project timeline. Criteria for selecting candidates should include demonstrable fluency with the game engine to be used in the project and a portfolio demonstrating past work. At the outset of the project under discussion, a call for portfolios was sent out to the entire game design community on and off campus. However, even though the positions were paid, only two students submitted portfolios. One asked if the game produced for the grant could also apply to his thesis project. One option for departments with graduate game design programs would be to consider building such projects into thesis classes each term so that the student's final deliverable would be a game for their portfolio. Alternatively, the team recommends relying on faculty recommendations for qualified candidates and reaching out to them individually and directly to solicit participation.

The production phase progressed quickly and smoothly. The recommendations for repeating such success follow the process outlined above. The timeline was appropriate for the group given the scale of the work. However, the scope of the project, number of levels, original assets to be created, etc. need be considered to establish a realistic development timeline for other projects. The grant budgeted 30 hours for the project manager, 20 hours for each game designer, and 35 hours for the coder. While the graduate students in the role of game designer found the time allotment to be ideal, they did note that would not be the case for a full-time student or those with other responsibilities. Also, the project manager and coder went over their allotted hours, thus the team recommends budgeting more by percentage for these roles.

With regards to the milestones and check in timeline, the group repeatedly noted how productive these were in terms of providing small goals to reach each week and getting regular feedback. Additionally, the regular meetings build rapport and camaraderie among the group, which leads to more effective teamwork. The Discord provided the opportunity to ask small, technical questions and have them answered quickly on an as-needed-basis to keep the project moving ahead. The project manager sent out weekly assignments, breaking down the expectations for each team member and providing a clear timeline and goals. The project manager also met with the team outside of weekly meetings to address any outstanding technical issues, such as advanced animations and character rigging. Other communication software and strategies may also be considered:

- Trello—tasks assigned per sprint review
- Slack—feedback on art process, general announcements, DMs on specific questions
- Teams—video conference sprint reviews
- Google Drive—storage of assets, archive of presentations, additional files
- Google Slides—sprint review template

The post-production phase was truncated and could be expanded. Working under the constraints of the hours identified for the internal grant, the schedule for semester end dates, and availability of playtesters, the team agreed that more time should have been allotted for this phase. With the playtesting taking place over the course of two weeks, there was insufficient time to address all of the recommendations from respondents of the survey. Finally, sharing files on Google Drive did prove to be sufficient for the scope of the project, but having multiple versions that up to five individuals were working on that needed to be merged and updated continuously was laborious. Project manager recommends moving to Git Hub or using other version control software.

5. Recommendations

5.1. Initial Considerations

The first consideration for any organization (or individual) planning to create an educational experience would be to assess capabilities and capacities at hand. Resources and support available will determine the size and scope of the project that may be undertaken. In many instances the SME themselves may have skills in coding or game design and development that would enable the completion of a small project or minimum viable product (MVP) that can be used to garner more funding. However, for large immersive experiences with multiple levels, assessments, and UI elements, a team will be necessary, especially since the goal of an educational game studio would be to afford the creation of experiences from a range of disciplines. At the same time, the team need not be solely comprised of individuals from within the organization: partnerships and an interorganizational network should be cultivated to maximize skills and resources. For instance, one may have subject matter expertise with researchers and scholars on a given topic and another technical expertise and personnel to develop a proposed GDD.

5.2. Structural Considerations

Regardless of whether an institution has the subject matter expertise or a game design department with technical expertise to realize educational gaming experiences, some guidelines for team and studio structure should be considered. For instance, results from the study confirm the necessity of having a clear project management plan supported by multiple avenues for regular communication and key roles in place. The size and scope of the project will determine the best path forward with regards to team structure, communication strategies, and incentives for participants. As noted, if the project involves a single-scene cinematic, VR mapping of one location, simple immersive simulation or single-level video game, then a small team will suffice depending upon their set, whereas large, multi-level games, full cinematic or multiple-level or location VR mapping will require a larger team. The timeline for team springs should reflect the size and duration of the project. For instance, if a project is to be completed within a single semester (between 3-4 months), weekly check-ins for the team to connect and report are advisable. On the other hand, if a project is slated for two semesters (9 months or longer), sprints may be stretched to 2 - 3 weeks. The following suggested structural models are designed to accommodate mid- to large-sized gamification projects.

Option 1: *Internal Game Studio*

The first option would be to design a studio that is funded either by an internal endowment or through experiential learning credit. Regardless, whenever possible, the project manager should consider compensating students, especially undergraduates, to ensure a high caliber of work produced. Funding for student and other team members may also be provided through an endowed fund that is renewable annually and used for internal projects for research in general or specifically for gamification projects.

The team structure need be designed to accommodate the needs of those involved: whether the team is local or will work virtually, how much and what variety expertise throughout the development process will be needed, and what the timeline will be for the final deliverable. In the internal model, a faculty member will need a course release(s) or additional compensation to oversee management of the teams and projects, or an organization may hire a full-time staff member or post-doctoral student for a fellowship.

Option 2: *Collaborative Game Studio*

The second option would be to have an educational institution partner with another external entity, either a for-profit studio or another institution that has a game design department. In the external partnership model, students, faculty, and staff would still be involved in the creation of projects. The approach would be best suited for those who have SME but not a game design program proper at their own organization. As such, a faculty member with expertise in an area with the complete GDD would work with students in that field to provide the instructional materials (audio recordings, quizzes, etc.) and oversight. In this common instance, the external partner would be providing technical expertise and, if an educational institution, experience for their own students in game design or computer science.

Option 3: *Hybrid Game Studio*

The hybrid model can be staffed by internal personnel or through the collaborative model. In addition to creating bespoke educational experiences for use in higher education classes, this hybrid approach also gives students real-world experience in working with clients in the for-profit sector. Establishing an LLC will be required to operate and handle revenue generation outside of the educational non-profit sector.

5.3. Roles in Teams

There are many variables that need to be considered when identifying who should be on a project team. In deciding on these roles, some questions should be answered that also affect the timeline and budget.

- Does the SME already have all the research required for the project completed and instructional materials designed?
  - If not, consider extending the timeline and including research assistants to budget
- Is the project being created using hardware and software familiar to the team?
  - If not, consider including a collaborative model or contractor to fill gaps in knowledge or include in timeline and budget to allow for upskilling team
- Are there specific tasks to be completed that require specialized skills or knowledge?
  - Identify these tasks and the associated expertise needed
  - Consider budgeting average hours per week per role *(note*: an individual may fulfil multiple roles, thus alleviating the duplication of costs for the project)
    - Game Director (40+ hours/week)
    - Art Director (40+ hours/week)
    - Designer (40+ hours/week)
    - Programmer (40+ hours/week)
    - Sound Design (40+ hours/week)
    - Tech Art (40+ hours/week)
    - UI/UX Designer (40+ hours/week)
    - 3D Character Modeler (20+ hours/week)
    - Assistant Programmer (20+ hours/week)
    - Character Texture (20+ hours/week)
    - Rigging (20+ hours/week)
    - Concept Artist (5 - 8 hours/week)
    - Storyboard Artist (5 - 8 hours/week)

- 2D Textures Artist (5 - 8 hours/week)
- 2D Cinematics (15+ hours/week)
- 3D Character Animation/Rigging (15+ hours/week)
- 3D Modeler/Environment Artist (20+ hours/week)

5.4. Role of the SME

The role of the subject matter expert (SME) may shape the roles of the team, as well. The SME may be internal or external to the institution; their relationship with the project may be direct or indirect. Therefore, the following roles are proposed for the SME that range from most to least involved.

- *Project Manager*: In this role, the SME provides both subject matter expertise in a given area for the topic to be gamified and possesses a working knowledge of game design and production. The team structure is best suited to smaller projects whereby a small team may be recruited to support development. The SME may also serve in the "two-boss system" as a project manager.
- *Team Member*: In this role, the SME serves as dedicated member of development team throughout process after completing the initial GDD and/or LDD, depending upon scope of project. The team is led by a project manager with expertise in design and development, but the project may be complex and evolving, requiring the SME to remain in the design spring meetings to answer questions and trouble shoot.
- *Subject Expert*: In this role, the SME provides the initial premise for game and complete description of concept and mechanics via GDD, but (much like in the instructional design process) hands off the materials to the development team and is contacted on an as needed basis should clarification on specific elements arise. The success of this role relies upon a thorough Game Design Document template being provided and consulting meetings completed with a game designer/developer to ensure all information is known from the outset and a fully realized experience may be completed separate from the original author.

5.5. Financial and Curricular Considerations

In all the options outlined above, the work performed by students can be embedded within curricular, co-curricular or extracurricular elements. The appropriate selection will depend upon the option selected above and whether the curricular offerings exist, such as with a game design and development or computer science program that uses game engines. As such, Option 1 would lend itself well to embedding projects within the game design curriculum proper. In instances where there may not be such a program, such as with Option 2, offering experiential credit for students across a range of disciplines is a better option. Finally, Option 3 may be used with co-curricular or extracurricular considerations, such as direct compensation for work on projects.

6. Conclusions

Emerging technologies continue to disrupt higher education models. Educators are increasingly exposed to generations of learners that expect engagement and participatory experiences. One solution outlined in this study is to take advantage of existing resources within and external to institutions. The software and game engines that undergird the creation of gamified learning experiences are free for educational use. The asset stores for the major engines, including Unity and Unreal Engine, boast thousands of free assets to populate and may be recombined for an unlimited number of scenarios and learning experiences in different fields. Where those interested in pursuing gamification of learning run aground is in finding those with the expertise or the time to upskill themselves or their students to take advantage of these resources. With that in mind, the results of the study encourage administrators, IT, and centers for teaching excellence to hold workshops and provide free training to faculty and staff. Incentivizing upskilling with professional development recognition, modest compensation, and/or certificates awarded via learning academies will improve buy in from constituents. However, support from the administration is key—an institution must make the strategy part of their institutional goals and put resources behind the initiative. Future research is recommended to scale out the designs outlined in here and determine how additional efficiencies may be leveraged in a given institutional structure.

Institutions should remain relevant in the coming decade. The new participatory culture demanded by incoming students needs be recognized through engaging instructional material. Gone are the days of reading a textbook, listening to a lecture in class, and then taking an exam. In the information age, and as digital natives, students have near limitless access to information at their fingertips in real time and on demand. The role of educators will increasingly be to step aside and facilitate active learning strategies, stepping in, not as the sage on the stage, but as the SME who can assist with developing durable skills, information literacy, and higher-order thinking to apply knowledge in various contexts. The gamification of education is one such tool at the disposal of educators today to assist in this way and ensure that immersive learning experiences are meaningful, impactful, and engaging.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Papadakis, S. and Kalogiannakis, M. (2017) Mobile Educational Applications for Children: What Educators and Parents Need to Know. *International Journal of Mobile Learning and Organisation*, **11**, 256-277. https://doi.org/10.1504/IJMLO.2017.085338

[2] Hossein-Mohand, H., Trujillo-Torres, J.M., Gómez-García, M., Hossein-Mohand, H. and Campos-Soto, A. (2021) Analysis of the Use and Integration of the Flipped Learning Model, Project-Based Learning, and Gamification Methodologies by Secondary School Mathematics Teachers. *Sustainability*, **13**, Article No. 2606. https://doi.org/10.3390/su13052606

[3] Dastyar, S. (2019) The Investigation of the Effectiveness of Participatory Learning Education on Students Motivation and Academic Achievement. *International Journal of Advanced Research and Publications*, **3**, 165-170.

[4] Squire, K. (2011) Video Games and Learning: Teaching and Participatory Culture in the Digital Age. Columbia University, New York.

[5] Adams, S.P. and Du Preez, R. (2022) Supporting Student Engagement through the Gamification of Learning Activities: A Design-Based Research Approach. *Technology, Knowledge and Learning*, **27**, 119-138. https://doi.org/10.1007/s10758-021-09500-x

[6] Bundrage, C. and Mapson, K. (2022) Design and Development of an Online Professional Development Course on Culturally Responsive Pedagogy Using the ADDIE Model. In: *Society for Information Technology & Teacher Education International Conference*, Association for the Advancement of Computing in Education (AACE), Morgantown, 248-256.

[7] Kennette, L.N. and Beechler, M.P. (2019) Gamifying the Classroom. *Transformative Dialogues: Teaching and Learning Journal*, **12**, 1-8.

[8] Deterding, S., Sicart, M., Nacke, L., O'Hara, K. and Dixon, D. (2011) Gamification: Using Game Design Elements in Non-Gaming Contexts. *Proceedings of the* 2011 *Annual Conference Extended Abstracts on Human Factors in Computing Systems*, Vancouver, 7-12 May 2011, 2425-2428. https://doi.org/10.1145/1979742.1979575

[9] Reynaldo, C., Christian, R., Hosea, H. and Gunawan, A.A. (2021) Using Video Games to Improve Capabilities in Decision Making and Cognitive Skill: A Literature Review. *Procedia Computer Science*, **179**, 211-221. https://doi.org/10.1016/j.procs.2020.12.027

[10] Granic, I., Lobel, A. and Engels, R. (2014) The Benefits of Playing Video Games. *American Psychologist*, **69**, 66-78. https://doi.org/10.1037/a0034857

[11] Al-Moteri, M., Alrehaili, A.A., Plummer, V., Yaseen, R.W., Alhakami, R.A., Al Thobaity, A. and Faizo, N.L. (2021) Rapid Visual Search Games and Accuracy of Students' Clinical Observation Skills: A Comparative Study. *Clinical Simulation in Nursing*, **55**, 19-26. https://doi.org/10.1016/j.ecns.2021.03.003

[12] Caroux, L. and Mouginé, A. (2022) Influence of Visual Background Complexity and Task Difficulty on Action Video Game Players' Performance. *Entertainment Computing*, **41**, Article ID: 100471. https://doi.org/10.1016/j.entcom.2021.100471

[13] Li, J., Zhou, Y. and Gao, X. (2022) The Advantage for Action Video Game Players in Eye Movement Behavior during Visual Search Tasks. *Current Psychology*, 1-10. https://doi.org/10.1007/s12144-022-03017-x

[14] Laamarti, F., Eid, M. and El Saddik, A. (2014) An Overview of Serious Games. *International Journal of Computer Games Technology*, **2014**, Article ID: 358152. https://doi.org/10.1155/2014/358152

[15] Dörner, R., Göbel, S., Effelsberg, W. and Wiemeyer, J. (2016) Serious Games. Springer, Cham. https://doi.org/10.1007/978-3-319-40612-1

[16] Hutson, J. and Fulcher, B. (2022) A Virtual Reality Educational Game for the Ethics of Cultural Heritage Repatriation. *Games and Culture*, 1-19. https://doi.org/10.1177/15554120221131724

[17] Kim, S., Song, K., Lockee, B. and Burton, J. (2018) Gamification in Learning and Education: Enjoy Learning Like Gaming. Springer, Cham. https://doi.org/10.1007/978-3-319-47283-6

[18] Wideman, H.H., Owston, R.D., Brown, C., Kushniruk, A., Ho, F. and Pitts, K.C. (2007) Unpacking the Potential of Educational Gaming: A New Tool for Gaming Research. *Simulation & Gaming*, **38**, 10-30. https://doi.org/10.1177/1046878106297650

[19] Krath, J., Schürmann, L. and Von Korflesch, H.F. (2021) Revealing the Theoretical Basis of Gamification: A Systematic Review and Analysis of Theory in Research on Gamification, Serious Games and Game-Based Learning. *Computers in Human Behavior*, **125**, Article ID: 106963. https://doi.org/10.1016/j.chb.2021.106963

[20] Sailer, M., Hense, J., Mandl, H. and Klevers, M. (2017) Fostering Development of Work Competencies and Motivation via Gamification. In: Mulder, M., Ed., *Competence-Based Vocational and Professional Education*, Springer, Cham, 795-818. https://doi.org/10.1007/978-3-319-41713-4_37

[21] Adams, N., Little, T.D. and Ryan, R.M. (2017) Self-Determination Theory. In: Wehmeyer, M.L., *et al*., Eds., *Development of Self-Determination through the Life-Course*, Springer, Dordrecht, 47-54. https://doi.org/10.1007/978-94-024-1042-6_4

[22] Deci, E.L., Koestner, R. and Ryan, R.M. (1999) A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, **125**, 627-668. https://doi.org/10.1037/0033-2909.125.6.627

[23] Dweck, C.S. and Leggett, E.L. (1988) A Social-Cognitive Approach to Motivation and Personality. *Psychological Review*, **95**, 256-273. https://doi.org/10.1037/0033-295X.95.2.256

[24] Elliott, E.S. and Dweck, C.S. (1988) Goals: An Approach to Motivation and Achievement. *Journal of Personality and Social Psychology*, **54**, 5-12. https://doi.org/10.1037/0022-3514.54.1.5

[25] Nicholls, J.G. (1984) Achievement Motivation: Conceptions of Ability, Subjective Experience, Task Choice, and Performance. *Psychological Review*, **91**, 328-346. https://doi.org/10.1037/0033-295X.91.3.328

[26] Hamstra, M.R.W., van Yperen, N.W., Wisse, B. and Sassenberg, K. (2014) Transformational and Transactional Leadership and Followers' Achievement Goals. *Journal of Business and Psychology*, **29**, 413-425. https://doi.org/10.1007/s10869-013-9322-9

[27] Pekrun, R., Cusack, A., Murayama, K., Elliot, A.J. and Thomas, K. (2014) The Power of Anticipated Feedback: Effects on Students' Achievement Goals and Achievement Emotions. *Learning and Instruction*, **29**, 115-124. https://doi.org/10.1016/j.learninstruc.2013.09.002

[28] Seifert, T. (2004) Understanding Student Motivation. *Educational Research*, **46**, 137-149. https://doi.org/10.1080/0013188042000222421

[29] Bandura, A. (1977) Self-Efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, **84**, 191-215. https://doi.org/10.1037/0033-295X.84.2.191

[30] Perron, B. and Wolf, M.J. (2009) The Video Game Theory Reader 2 (Vol. 2). Routledge, New York. https://doi.org/10.4324/9780203887660

[31] Callele, D., Neufeld, E. and Schneider, K. (2005) Requirements Engineering and the Creative Process in the Video Game Industry. 13*th IEEE International Conference on Requirements Engineering*, Paris, 29 August-2 September 2005, 240-250. https://doi.org/10.1109/RE.2005.58

[32] Chandler, H.M. (2009) The Game Production Handbook. Jones & Bartlett Publish-

ers, Burlington.

[33] Weitze, C.L. (2021) Recommendations for Learning through Educational Game Design: A Systematic Literature Review. In: *European Conference on Games Based Learning*, Academic Conferences International Limited, Reading, UK, 745-XII.

[34] Politowski, C., Petrillo, F., Ullmann, G.C. and Guéhéneuc, Y.G. (2021) Game Industry Problems: An Extensive Analysis of the Gray Literature. *Information and Software Technology*, **134**, Article ID: 106538. https://doi.org/10.1016/j.infsof.2021.106538

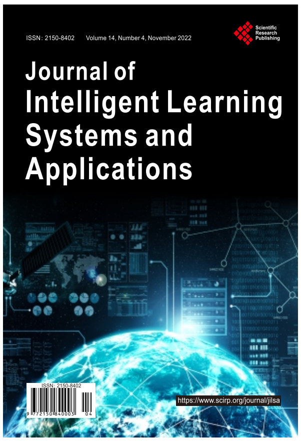
[35] Shave, L. (2022) Executive Management and Teams: A Hybrid Model for Workplace, Remote Working and Virtual Team Management. *IQ*: *The RIMPA Quarterly Magazine*, **38**, 20-24.

[36] Garro-Abarca, V., Palos-Sanchez, P. and Aguayo-Camacho, M. (2021) Virtual Teams in Times of Pandemic: Factors That Influence Performance. *Frontiers in Psychology*, **12**, Article ID: 624637. https://doi.org/10.3389/fpsyg.2021.624637

[37] Colby, R. and Colby, R.S. (2019) Game Design Documentation: Four Perspectives from Independent Game Studios. *Communication Design Quarterly Review*, **7**, 5-15. https://doi.org/10.1145/3321388.3321389

[38] Dimitriadou, A., Djafarova, N., Turetken, O., Verkuyl, M. and Ferworn, A. (2021) Challenges in Serious Game Design and Development: Educators' Experiences. *Simulation & Gaming*, **52**, 132-152. https://doi.org/10.1177/1046878120944197

[39] Winget, M.A. and Sampson, W.W. (2011) Game Development Documentation and Institutional Collection Development Policy. *Proceedings of the* 11*th Annual International ACM*/*IEEE Joint Conference on Digital Libraries*, Ottawa, 13-17 June 2011, 29-38. https://doi.org/10.1145/1998076.1998083



Call for Papers

Journal of Intelligent Learning Systems and Applications

ISSN Print: 2150-8402 ISSN Online: 2150-8410
https://www.scirp.org/journal/jilsa

The ***Journal of Intelligent Learning Systems and Applications*** **(JILSA)** is a peer reviewed international journal with a key objective to provide the academic and industrial community a medium for presenting original cutting-edge research related to intelligent learning systems and their applications. JILSA invites authors to submit their original and unpublished work that communicates current research on intelligent learning systems both in the theoretical and methodological aspects, as well as various applications in real-world applications.

Papers are invited on the topics including, but not limited to:

- Approximate Dynamic Programming
- Autonomic Computing
- Autonomous Learning Systems
- Bio-Inspired Learning Method
- Clustering
- Cyber-Physical Systems
- Evolutionary Computation
- Filter Bank
- Fuzzy and Rough Set
- General Theory on Intelligent Learning Systems
- Intelligent Applications
- Learning Control Systems
- Multi-Agent Learning
- Multiple Models Strategy
- Neural Networks
- Pattern Recognition Based on Learning Techniques
- Predictive Control
- Recommendation System
- Reinforcement Learning
- Robotics
- Smart Grid Techniques and Applications
- Social Network Analysis
- Statistical Learning Theory
- Supervised Learning
- Time-Frequency Representation
- Unsupervised Learning
- User Behavior Modelling
- Wavelet Theory

Editor-in-Chief

Prof. Wei-Chiang Hong — Asia Eastern University of Science and Technology, Chinese Taipei

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Website and E-Mail

https://www.scirp.org/journal/jilsa **E-Mail: jilsa@scirp.org**