22nd International Conference on Advances in ICT for Emerging Regions (ICTer 2022): 019 - 024

Identification of Previously Unseen Asian Elephants using Visual Data and Semi-Supervised Learning

Gagana Weerasinghe[1], Kasun Karunanayaka[1], Prabhash Kumarasinghe[1], Dushani Perera[1], Carlos Trenado[2], Kasun De Zoysa[1], Chamath Keppitiyagama[1],

[1]*University of Colombo School of Computing, Sri Lanka and*

[2]*Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany*

2017cs192@stu.ucsc.cmb.ac.lk, ktk@ucsc.cmb.ac.lk, jpk@ucsc.cmb.ac.lk, udi@ucsc.cmb.ac.lk, carlos.trenadoc@gmail.com, kasun@ucsc.cmb.ac.lk, cik@ucsc.cmb.ac.lk

2022 22nd International Conference on Advances in ICT for Emerging Regions (ICTer) | 979-8-3503-4613-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICTer58063.2022.10024068

Abstract*— This paper presents a novel method to identify unseen Asian elephants that are not previously captured or identified in available data sets and re-identify previously seen Asian elephants using images of elephant ears, leveraging a semi-supervised learning approach. Ear patterns of unseen elephants are learnt for future re-identification. To aid our process, elephant ear patterns are used as a biomarker to uniquely identify individual Asian elephant, each of which is attached a descriptor. The main challenge is to learn and use a clustering technique to identify new classes (i.e., elephants) in unlabelled elephant ear image sets and leveraging this data in verifying the labelled images. This study proposes a systematic approach to address the problem to uniquely identify elephants, where we developed: (a) a self-supervised learning approach for training the representation of labelled and unlabelled image data to avoid unwanted, bias labelled data, (b) rank statistics for transferring the models' knowledge of the labelled classes when clustering the unlabelled images, and, (c) improving the identification accuracy of both the classification and clustering algorithms by introducing a optimization problem when training with the data representation on the labelled and unlabelled image data sets. This approach was evaluated on seen (labelled) and unseen (unlabelled) elephants, where we achieved a significant accuracy of 86.89% with an NMI (Normalized Mutual Information) score of 0.9132 on identifying seen elephants. Similarly, an accuracy of 54.29% with an NMI score of 0.6250 was achieved on identifying unseen elephants from the unlabelled Asian elephant ear image data set. Findings of this research provides the ability to accurately identify elephants without having expert knowledge on the field. Our method can be used to uniquely identify elephants from their herds and then use it to track their travel patterns which is greatly applicable in understanding the social organization of elephant herds, individual behavioural patterns, and estimating demographic parameters as a measure to reducing the human-elephant conflict in Sri Lanka.

Keywords*— *Asian Elephants, Biomarkers, Semi-Supervised learning, Clustering, Identification, Visual descriptor

I. Introduction

Human have been fascinated by elephants for thousands of years as it is the largest living land mammal in the world [1]. These creatures are adored by people all over the world and are main attractions in tourism, helping to draw funds that can be used to protect wildlife. There are primarily two types of elephants: (a) African (*Loxodonta africana*), and (b) Asian elephants (*Elephas maximus*) [2]. Over the past 75 years, Asian elephant population has decreased by 50% and only an approximated amount of 20,000 to 40,000 are left in the nature [3]. Currently, the wild elephant population in Sri Lanka is about 7,500 and 361 out of that have died during 2019 [4]. The International Union for the Conservation of Nature (IUCN) has considered the Sri Lankan elephant to be endangered [3]. The situation is to get worse in the next few years due to the human-elephant conflict and the rate of habitat loss [5].

Photographic elephant identification has been the most widely used method for identification and marking elephants. This is done by comparing physical characteristics and finding mismatches between the new elephant images with existing images in the database such as tail, back, tush/task, ear [6]. This is however a complex task and ecologists spend many hours in charting and cataloguing features to distinguish elephants individually. This human-centric operation is highly subjective because it depends on the observer's skills and interpretation. It is therefore important to automate these existing human-made methods for large-scale analysis of these species for research, conservation, and tracking. With the advancement of computer vision, image recognition and pattern recognition techniques, identifying elephants can be more accurately performed, reducing the time-consumed.

There have only been a few attempts made to automate identifying elephants using machine learning and computer vision. Ardovini et al. [7], Korschens, and Denzler [8], Dabarera and Rodrigo [9] and Kulits [10] identify elephants using a visual technique which has been tested on African elephants but not on Asian elephants. However, there has been no study done to properly identify unseen new elephants that were not in the training dataset, let alone using ear images of the elephants to learn new unseen patterns for the use of future re-identification. We also consider the whole ear pattern as a biomarker for Asian elephants to identify and re-identify (verification). In previous work, re-identification is performed only leveraging supervised approaches that utilize training datasets and discovering new unseen Asian elephants is yet to be investigated. Using semi-supervised learning to identify elephants will improve the current level of accuracy and also provides the potential to learn new patterns.

The following research questions are addressed in this paper:

1. How to visually identify a new Asian elephant that is not in the training set at an individual level?

2. How to assign a visual descriptor per ear of an Asian elephant?

The following contributions are made in this paper:

- we design and develop a semi-supervised learning model for identification and discovering of Asian elephants uniquely using ear images,
- we develop and assign visual descriptors for elephant ears,
- we develop an algorithm to identify and tag new elephants that are not previously identified in the datasets by learning new ear patterns of new unseen elephants.

Rest of the paper is structured as follows. In Section II, we discuss the related work in the area. Section III describes our semi-supervised technique to identify known elephants and learn ear patterns of previously unseen elephants. Section IV presents the results obtained using our technique. Finally, we make conclusions and provide future direction to the research area in Section V.

II. Related work

In this section we critically review previous work in identifying elephants using computational techniques such as biometrics identification, deep learning methods, visual feature descriptors, contrastive learning, etc.

A. *Elephant Re-identification*

Sangineto et al. [7] identifying the nick shape patterns of African elephants using a semi-automated elephant photo identification technique utilizing multi-curve matching. This study has considered only one feature – the ear's nick. Check Position algorithm and Global Matching algorithms are responsible for matched nicks with a corresponding valid global position. But this nick feature is common on African elephants, and it is not very common in Asian elephants.

Korschens et al. [8] also presented a matching algorithm formed on human-labelled whole-head interpretation which includes elephant's tusks and ears. They also follow the same method for ears as in [7] which matches the nick characteristics of the elephants. A system using YOLO object detector as baseline has been trained on discovering elephant heads while ImageNet has been trained with ResNet50 feature extraction. Elephant classification has been performed using a Support Vector Machine. Authors present an accuracy of 80% on top-10 and 56% on top-1 accuracy for African elephants.

Kulits et al. [10] have also proposed elephant re-identification model. They proposed two types of models: (a) extracting features of ear contours from images based on Weideman et al. [11] contour matching algorithm, and (b) the hybrid model which uses the contour matching and a SEEK algorithm. In order to slender the set of possible matches that should be examined by human experts to allow curvature matching algorithm to focus on the ear, they trained a small elephant-ear detector which was based on a Faster R-CNN object detection model with a ResNet-50 backbone with the addition of Feature Pyramid Networks. Authors have used the ELPephants dataset which was introduced and created by Korschens et al. [8].

De Silva et al. [12] proposed a model for elephant re-identification on Asian elephants based on ears, face, and full body as features. They presented a CNN based approach to identify Asian elephants consisting of several steps: (a) elephant-ear localisation step at a species-level, and (b) an ear-patch classification step at an individual-level. For detection of the ears of an elephant they used YOLO which helps localization of the ears using four datasets. Authors compare five types of CNN models for classification: Alexnet, VGG16, ResNet50 , InceptionV3 and Xception. Xception model with the ele-ear dataset has outperformed other combinations with a top-1 accuracy of 88% and top-5 accuracy of 99.27%. This in turn gives the conclusion that elephant ears contain most features which can be considered as the good biomarker for recognition and gives the best accuracy with the Xception model.

B. *Deep Learning Methods for Animal Re-Identification*

In the previous studies there are various research in animal re-identification using deep learning techniques. In [15], the spine of cows has been used as the feature while in [13] and [14], fins of dolphins and minke whales were used. Authors in [13] and [15] have used Softmax classification to assign identities using the features extracted from the CNN.

There are various studies which have used facial recognition such as in [18] (Gorilla), [17] (Chimpanzees), [16] (Red Pandas), and [19] (Cow). These models first detect and localize the animal in the picture using an object detector and then use that localized part for feature extraction. The mentioned studies are also based on CNN with AlexNet, VGG-16, InceptionV3, ResNet as the backbone network.

C. *Open-world Learning and Novel discovering*

Open-world learning and novel class discovery has been discussed by Kaidi Cao et al. [21]. Authors develop a method called ORCA (Open-worRld with unCertainty based Adaptive margin) which helps to discover new unseen classes, as well as identify previously known classes using an end-to-end deep learning framework. This method clusters and classifies data simultaneously. There will be additional classification heads for the number of anticipated unseen classes and a linear classifier with heads for seen classes. ORCA uses pairwise objective and supervised objective to eventually create pseudo-labels for unlabelled data. Feature representation for both unlabelled and labelled data is done by first applying the embedding function pre-trained using self-supervised learning. Standard benchmark classification of image data sets with both labelled and unlabelled data have been used to evaluate ORCA: ImageNet, CIFAR-100 and CIFAR-10. Authors have achieved an accuracy of 89.1% for seen classes and 72.1% for novel classes in the ImageNet-100 dataset.

In the domain of novel class identification, Kai Han et al. [22] presented a study on automatically discovering and learning new visual categories with ranking statistics. They have solved the issue by finding novel class photo collections in other labelled classes. Three approaches have been combined to address this issue: (a) using a self-supervised learning using the RotNet to bootstrapping a picture representation, (b) rank statistics used to send knowledge of the labelled class in the model to the problem of unlabelled pictures and clustering, and (c) supervised classification for labelled and clustering unlabelled data improved by optimizing the joint objective function on both labelled and unlabelled data. They evaluated the model with CIFAR-10, CIFAR-100, SVHN, OmniGlot and ImageNet. They have used the evaluation metrics Clustering accuracy (ACC) to evaluate clustering performance. They have achieved an accuracy of 91.7%, 75.2% and 95.2% for CIFAR-10, CIFAR-100 and SVHN with incremental learning.

III. Semi-Supervised learning approach to identify elephants

Our proposed model is based on the automatic discovery and learning new visual categories with ranking statistics [22]. We were able to tackle the problem of reidentifying the labelled data (seen elephants) and identifying the new classes (new unseen elephants) in unlabelled data using this model, which was a clustering problem. In the initial model it was based on three main steps as depicted in Fig.1:

1. First, we pre-train the image representation (using a CNN) for both labelled and unlabelled data using a self-supervised learning objective,
2. Then, we learn the higher-level features of the labelled data using a fully supervised method, and
3. Finally, fine-tuned representation is used, via rank statistics, to induce clusters in the unlabelled data, while maintaining a good representation on the labelled set.

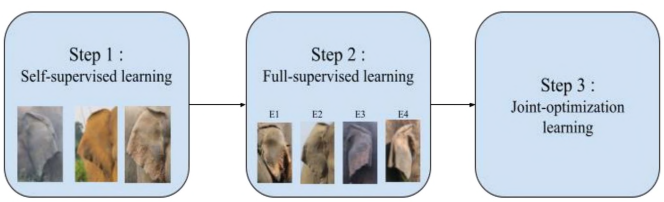

Figure 1: Proposed AutoNovel model design

A. Self-Supervised learning

Self-supervised learning is used to learn the features and the model without using labels for all the data. For this purpose, we have learnt all the labelled data (seen elephants) and unlabelled data (unseen elephants) without considering any labels. We use SimCLR model [23] for this purpose. SimCLR is a simple framework for contrastive learning of visual representations. SimCLR increases agreement through latent space's contrastive loss among data in the same data example that are of different augmented views, in order to learn representations. SimCLR consists of 4 components.

- Resulting two correlated views by transforming any data example using a stochastic data augmentation module,
- ResNet architecture used to represent vectors to be extracted from augmented data examples,
- Mapping and projection component that use a multilayer perceptron (MLP) consists of one hidden layer where contrastive loss is applied, and
- A component that performs a contrastive prediction task defined using a contrastive loss function.

For our model we use three small augmentations that have been applied sequentially: (a) random Gaussian blur, (b) random colour distortions, and (c) random cropping. This self-supervised learning step is applied to both labelled and unlabelled data. This is trained using the ResNet-50 architecture CNN model.

B. Fully supervised learning

The proposed model consists of two classification heads, one is to re-identify the seen elephants and the other classification head to discover the new unseen elephants. Re-identification classification head is trained in a fully supervised manner using the labels for it.

In this step, the labelled data is trained using the labels and the self-supervised model data is used in this process as well. The SimCLR trained model for all the data is fine-tuned with seen elephant data set with the labels. With this process the Classifier 1 (η^l) is trained based on it.

This learning implementation is also based on a basic block CNN ResNet architecture. We have halved every set of 10 epochs with a stepwise decaying learning rate beginning from 0.05 and fine-tuned the model for the labelled set for 200 epochs. Fine-tuning was done on the last convolutional block along with the linear classifier while fixing the first three convolutional blocks of the model. As for the error function, we consider the Cross entropy (CE) loss function.

$$L_{\mathrm{CE}} = -\frac{1}{N}\sum_{i=1}^{N}\log \eta^{l}_{y_i}(z^{l}_{i}) \quad (1)$$

; where image x^l_i's self-supervised representation is given by z^l_i. Labelled data overfitting has been avoided by only updating the self-supervised model's last macro-block and η^l which is the classification head.

C. Joint-Optimization on Seen and unseen classes

In the next steps, a new classification head is trained and learnt to discover new novel classes in unseen elephants using the fine-tuned supervised model and with the self-supervised pseudo labelled dataset. It is now a clustering problem. We used the ramp-up function for the consistent regularization. The merged set of both unlabelled and labelled data has been sampled randomly with the joint training step. In this particular learning process, we have used two classification heads and loss functions respectively. We used the cross-entropy loss function as done in the supervised learning phase in the first classification head, where the descriptor is trained by optimizing the binary cross-entropy loss.

In this set, in order to define the relationships among them to aid the discovery process and to extract a descriptor vector and rank stat method for optimization purposes, a new classification head was learnt for the process of discovering unseen classes. The transfer learning process is used for discovering the unseen classes since their labels are unknown. The supervised trained model and self-supervised model are used for extracting the features and assigning the descriptor for unlabelled data. After the feature descriptor is extracted, we map the relation among two pairs of images with the extracted feature descriptor vector. We assume that two images are in the same class if the feature descriptors are proximal.

Rather than just comparing the descriptor vectors we use a rank statistic method to rank the vector values which is done using the magnitude. Two unlabelled images x^u_i and x^u_j belong to the same class if the rankings obtained for those are the same. Once this is done, we train unlabelled data using a comparison function with the help of the labels found earlier as pseudo-labels. Now we have two losses in hand related to the representation, both sharing the similar image embedding function which is self-supervised learning: the pairwise BCE loss L_{BCE} for the unlabelled data and the CE loss L_{CE} for the labelled data. We have also taken the Mean Squared Error (MSE) L_{MSE} as the consistency cost since we can consider this stage as a semi supervised learning process, because we are joining the labelled data classification and discovering process of unlabelled data.

$$L = L_{\text{CE}} + L_{\text{BCE}} + \omega(t) L_{\text{MSE}} \tag{2}$$

For the overall loss, coefficient ω(t) has been used as a ramp-up function. This model is also based on a CNN ResNet architecture. But in this model, after the first convolution layer it's followed by a Relu activation function with a max pooling layer. Finally, the two headers are assigned according to labelled and unlabelled

IV. Results and Discussions

For the testing and training our elephant identification technique, we used an elephant ear image dataset of 56 elephants containing 4100 ear images. Only the colour cropped ears of the elephants as shown in Fig. 2 are provided in the dataset. We split the data in two for training and testing on two categories as follow in Table I.

TABLE I. DATASET SUMMARY OF ELEPHANT EAR IMAGES

Description	Seen Elephants	Unseen elephants
Number of elephants	42	14
Number of images	3144	1048
Training images	2838	978
Testing images	206	70

Figure 2. Cropped Elephant ear images.

A. Self-Supervised learning (SimCLR)

First, we evaluated the self-supervised learning model. In this experiment, three sub experiments were conducted to investigate the impact of the number of epochs on test accuracy, i.e., 200, 500, 1000, respectively for training. For the validation of the model, a linear classifier was deployed where the model correctly classified 56 elephant classes with a top-1 accuracy of 33.33% (as shown in TABLE II).

TABLE II. RESULTS OF THE ELEPHANT EAR CLASSIFICATION MODEL USING SIMCLR WITHOUT LABLES.

Training epochs	Training Accuracy	Validation Accuracy
200	26.5 %	22.10 %
500	34.7 %	28.99 %
1000	39.9 %	33.33 %

Fig. 3 shows the class wise accuracy of our model with 1000 epochs for training. The self-supervised learning avoids the unwanted bias that is introduced by the frequent method of bootstrapping an image representation using the labelled data.

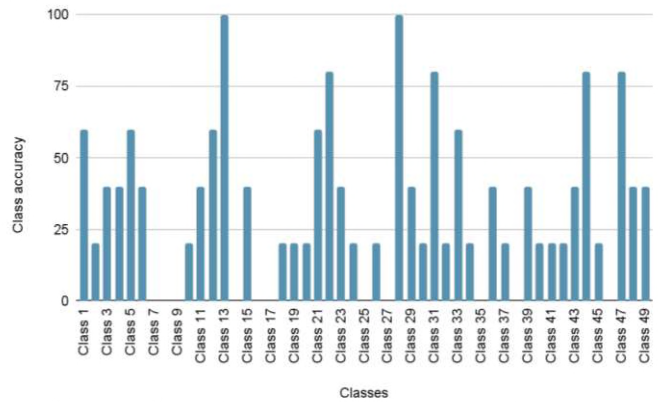

Figure 3. Class wise accuracy of SinCLR for 1000 epochs

B. Fully-Supervised learning for Seen elephant classes

Secondly, we evaluated the seen elephant classification process. We halved every set of 10 epochs with a stepwise decaying learning rate beginning from 0.05 and fine-tuned the model for the labelled set for 200 epochs. Ear images of 42 elephants (2838 images) were used for the model to train the supervised classification head which is based on a ResNet50 CNN model. Fig. 4 indicates the convergence of the model over 200 learning epochs.

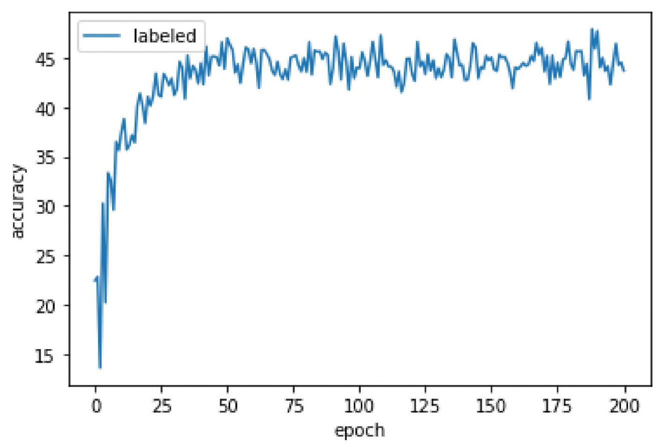

Figure 4. Training accuracy of fully supervised learning.

C. Joint optimization for re-identifying seen classes and novel class discovery for unseen classes

Thirdly, we evaluated the re-identification and novel elephant discovery process. The model is trained using 14 unseen elephants ear images (1048 images). Since this is performed as a joint optimization, the 42 seen elephant ear images (3144 images) will also be used in the optimization process. This model is also based on the ResNet50 CNN model. Training specifications are as follows:

- Epochs: 90
- Learning rate: 0.1
- Training images: 2838 & 978
- Testing images: 206 & 70

In order to evaluate the clustering performance, the clustering accuracy (ACC) is defined as follows:

$$\max_{g \in Sym(L)} \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}(\overline{y_i} = g(y_i)) \tag{3}$$

; where $\overline{y}_i$ and y_i give the clustering assignment and ground-truth label for every unlabelled data point x_i, L element group permutations are given by *Sym(L)*. Normalized Mutual Information score was also used to evaluate the clustering performance. NMI measures the agreement of the two assignments, ignoring permutations. These results are summarized in TABLE III and Fig.5 depicts the convergence of the model over 100 learning epochs.

TABLE III. ACCURACY, NMI SCORE AND MUTUAL INFORMATION SCORE FOR LABELED AND UNLABLED DATA.

Evaluator	Seen Elephants	Unseen Elephants
Accuracy	86.89 %	54.29 %
NMI score	0.9132	0.6250
Mutual Information	3.3907	1.6459

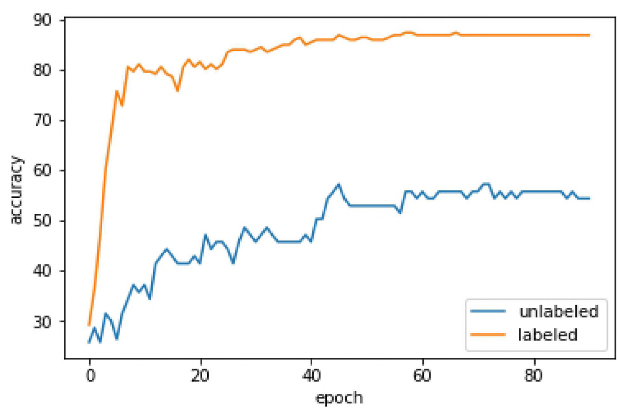

Figure 5. Accuracy for labelled and unlabelled data.

The model reported an 86.89% highest accuracy for identifying seen elephants and more importantly it achieved 54.29% accuracy of identifying unseen elephants from the unlabelled dataset. In Fig. 6 we visualize the t-distributed stochastic neighbour embedding (t-SNE) on the clustering of unseen elephants which were trained using the joint optimization. Each Newly identified elephant cluster is given a numerical number and a particular colour where each dot represents an elephant ear image. In Fig.9 it can be see that fourteen unseen elephants were clearly clustered except in a few classes.

V. CONCLUSIONS

This paper proposes a fully automated novel approach to discover unseen new elephants and identify previously seen elephants using elephant ears. Self-supervised learning was used in learning the representation of the elephant ears and the fine-tuned supervised learned model for seen elephants was used to identify and assign visual feature descriptors for elephant ears. With the aid of the feature descriptors and deep learning clustering, the proposed model achieved an accuracy of 86.89% identifying seen elephant classes and an accuracy of 54.29% discovering new unseen elephant classes which were not in the training dataset. We also proposed and evaluated another novel class identification model which was based on Progressive unsupervised learning (PUL). This uses a pre-trained model and fine tunes the model according to the new unseen unlabelled data by using k-means clustering. But after fine tuning the model, we discovered that the model was biased towards the new classes whereas the AutoNovel model could discover unseen classes to identify seen classes as well. The PUL model achieved a top-1 accuracy of 62.86% and a top-5 accuracy of 92.85%. This paper only focused on identifying and discovering Asian elephants. Therefore, input to the model must be an image of an elephant. This has been done based on the images where the elephant ear is clearly visible. The proposed solution could be further enhanced to increase the performance of the novelty discovering process and the overall design accuracy by suggested refinements. It could be further extended to identifying and discover of African elephants and any other animals as well.

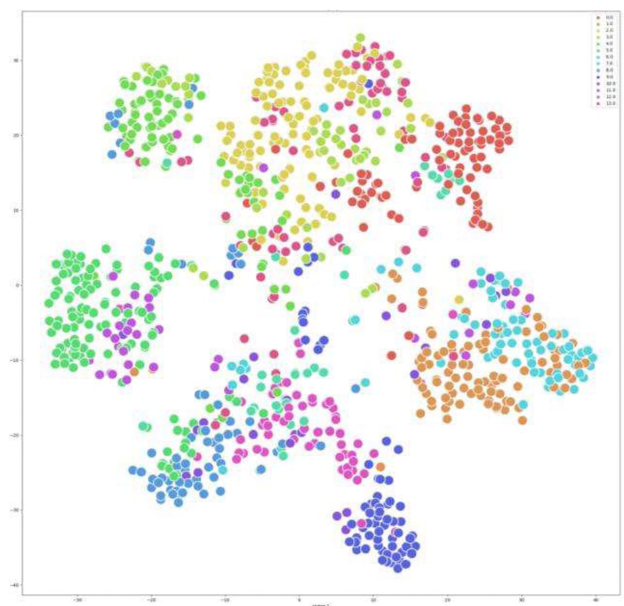

Figure 6: Evolution of the t-distributed stochastic neighbour embedding (t-SNE) on Unseen elephant data.

ACKNOWLEDGMENT

The first author would like to express his deepest gratitude to Dr. Carlos Trenado, Research Scientist at the Systems Neuroscience and Neurotechnology Unit in Saarland University, Germany for the support and assistance provided in carrying out this research. Further, the author would like to thank Mithsen De Silva for the assistance by providing the dataset of elephant images and useful insights for the improvement of the research. Further, the author like to express the sincere gratitude towards Dr. H.N.D Thilini, Dr. G.D.S.P Wimalaratne, and Dr. K.D Sandaruwan at the University of Colombo School of Computing for their immense support provided during the study. I would also like to thank Shakthi Weerasinghe for assistance with documentation and all our family and friends who helped, supported us in numerous ways to carry out this research successfully.

REFERENCES

[1] Riddle H, Schulte B A, Desai A A and Meer L: Elephants-A conservation overview Journal of threatened taxa (2010) 2 653-661

[2] Williams, C.; Tiwari, S.K.; Goswami, V.R.; de Silva, S.; Easa, P.S.; Kumar, A.; Baskaran, N.; Yoganand, K. & Menon, V. (2020). "Elephas maximus". IUCN Red List of Threatened Species.

[3] Asian Elephant: National geographic. Animals., 05-Apr-2021. from: https://www.nationalgeographic.com/animals/mammals/facts/asian-elephant

[4] BBC. (2020, January 11). Sri Lanka Elephants: 'record number' of deaths in 2019. BBC News, from: https://www.bbc.com/news/world-asia-51076898

[5] Hedges S and Gunaryadi D: Reducing human–elephant conflict: do chillies help deter elephants from entering crop fields? (2010) 44 139-146

[6] T. N. C. Vidya, D. Prasad, A. Ghosh, Individual identification in asianelephants, Gajah 40 (2014) 3–17.

[7] A. Ardovini, L. Cinque, E. Sangineto, Identifying elephant photos by multi-curve matching, Pattern Recognition 41 (2008) 1867–1877.

[8] M. Korschens, J. Denzler, Elpephants: A fine-grained dataset for elephant re-identification, in: 2019 IEEE/CVF International Conference on Computer Vision Workshops, ICCV Workshops 2019, Seoul, Korea (South), October 27-28, 2019, IEEE, 2019, pp. 263–270. URL:https://doi.org/10.1109/ICCVW.2019.00035. doi:10.1109/ICCVW.2019.00035.

[9] R. Dabarera, R Rodrigo. 2010. Vision Based Elephant Recognition for Management and Conservation. Fifth International Conference on Information and Automation for Sustainability (2010), 163–166. https://doi.org/10.1109/ICIAFS.2010.5715653

[10] Kulits, P., Wall, J., Bedetti, A., Henley, M., & Beery, S. (2021). ElephantBook: A Semi-Automated Human-in-the-Loop System for Elephant Re-Identification. ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS) (COMPASS '21), June 28-July 2, 2021, Virtual Event, Australia, 1(1). https://doi.org/10.1145/3460112.3472047

[11] Weideman, H. J., Stewart, C. V., Parham, J. R., Holmberg, J., J., & Pope, F. G. (2020). Extracting identifying contours for African elephants and humpback whales using a learned appearance model. Proceedings - 2020 IEEE Winter Conference on Applications of Computer Vision, WACV 2020, 1265–1274.

[12] De Silva, M., Kumarasinghe, P., De Zoysa, K., & Keppitiyagama, C. (Submitted). Reidentifying Asian Elephants from Ear Images Using a Cascade of Convolutional Neural Networks and Explaining with GradCAM. 1–14.

[13] D. A. Konovalov, S. Hillcoat, G. Williams, R. A. Birtles, N. Gardiner, M. I. Curnock, Individual minke whale recognition using deep learning convolutional neural networks, Journal of Geoscience and Environment Protection 6 (2018) 25–36. doi:https://doi.org/10.4236/gep.2018.65003.

[14] S. Bouma, M. D. M. Pawley, K. Hupman, A. Gilman, Individual common dolphin identification via metric embedding learning, in: 2018 International Conference on Image and Vision Computing New Zealand (IVCNZ), 2018, pp. 1–6. doi:https://doi.org/10.1109/ivcnz.2018.8634778.

[15] C. N. Phyo, T. T. Zin, H. Hama, I. Kobayashi, A hybrid rolling skew histogram-neural network approach to dairy cow identification system, in: 2018 International Conference on Image and Vision Computing New Zealand (IVCNZ), 2018, pp. 1–5

[16] Q. He, Q. Zhao, N. Liu, P. Chen, Z. Zhang, R. Hou, Distinguishing individual red pandas from their faces, in: Z. Lin, L. Wang, J. Yang, G. Shi, T. Tan, N. Zheng, X. Chen, Y. Zhang (Eds.), Pattern Recognition and Computer Vision, Springer International Publishing, Cham, 2019, pp. 714–724.

[17] A. Freytag, E. Rodner, M. Simon, A. Loos, H. S. K¨uhl, J. Denzler, Chimpanzee faces in the wild: Log-euclidean cnns for predicting identities and attributes of primates, in: B. Rosenhahn, B. Andres (Eds.), Pattern Recognition, Springer International Publishing, Cham, 2016, pp. 51–63. doi:https://doi.org/10.1007/978-3-319-45886-1_5.

[18] C. Brust, T. Burghardt, M. Groenenberg, C. Kading, H. S. K¨uhl, M. L. Manguette, J. Denzler, Towards automated visual monitoring of individual gorillas in the wild, in: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), 2017, pp. 2820–2830

[19] L. Bergamini, A. Porrello, A. C. Dondona, E. Del Negro, M. Mattioli, N. D'alterio, S. Calderara, Multi-views embedding for cattle re-identification, in: 2018 14th International Conference on SignalImage Technology Internet-Based Systems (SITIS), 2018, pp. 184–191. doi:https://doi.org/10.1109/sitis.2018.00036.

[20] Z. Li, C. Ge, S. Shen, X. Li, Cow individual identification based on convolutional neural network, in: Proceedings of the 2018 International Conference on Algorithms, Computing and Artificial Intelligence, ACAI 2018, Association for Computing Machinery, New York, NY, USA, 2018, pp. 1–5.

[21] Cao, K., Brbic, M., & Leskovec, J. (2021). Open-World Semi-Supervised Learning. 1–17. http://arxiv.org/abs/2102.03526

[22] Han, K., Rebuffi, S.-A., Ehrhardt, S., Vedaldi, A., & Zisserman, A. (2020). Automatically Discovering and Learning New Visual Categories with Ranking Statistics. 1–13. http://arxiv.org/abs/2002.05714

[23] Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. 37th International Conference on Machine Learning, ICML 2020, PartF168147-3, 1575–1585.

[24] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-January, 1800–1807. https://doi.org/10.1109/CVPR.