# Transfer Learning Based Myanmar Sign Language Recognition for Myanmar Consonants

Ni Htwe Aung[1], Ye Kyaw Thu[2], Su Su Maung[3], Swe Zin Moe[4], Hlaing Myat Nwe[5]

***Abstract*— In this paper, a study on the different Transfer Learning models is made for the purpose of recognizing Myanmar Fingerspelling (Myanmar Sign Language) alphabets. This experiment shows that Transfer Learning can play a significant role for sign language recognition system and is capable of recognizing the static hand gesture images that represent the Myanmar consonants from က (ka) to အ (a). The main objective of this paper is to investigate the performance of various Transfer Learning models for Myanmar Fingerspelling recognition. We proposed 12 Transfer Learning models using TensorFlow library and the accuracy for each model is compared. Among these 12 models, VGG16, ResNet50 and MobileNet with epoch 50 yielded the highest accuracy score with 94%. Although there are some limitations in the datasets, each model provides the encouraging results and thus, it can believe that the fully generalizable recognition system based on Transfer Learning can be produced for all Myanmar Sign Language Fingerspelling characters by doing further research with more data.**

***Index Terms*—Myanmar Sign Language, Myanmar Fingerspelling, Transfer Learning, Myanmar consonants.**

## I. Introduction

In recent years, some researchers have been paying attention to the research area of Sign Language (SL) recognition. It is important for many research fields as computer vision (CV), natural language processing (NLP), human computer interaction (HCI), image processing and computational linguistics. SL recognition system still remains as a challenging task because sign language is a visual language which contains the motion of the body, head, eyes, hands and facial expressions. SLs can differ from region to region and continent to continent based on the culture and environments of these particular regions. Therefore, it cannot be clearly said that how many SLs are used in the world. In Myanmar, there are 673,126 hearing-impaired persons according to the 2014 Myanmar national census [1]. Myanmar Sign Language (MSL) is mainly used by the Myanmar Deaf people to communicate with each other and other hearing people. MSLs used in southern Myanmar and northern Myanmar are also different. Moreover, there are very little research work in MSL recognition system. The proposed system would be the first transfer learning based MSL recognition system for Myanmar Fingerspelling consonants **'က'** (ka) to **'အ'** (a). This paper evaluated and investigated the accuracies of 12 different transfer learning models by using MSL images of Myanmar consonants that are currently using in southern part of Myanmar (mainly, teaching at the MaryChapman School for the Deaf, Yangon). We recorded the MSL videos of Myanmar consonants **'က'** (ka) to **'အ'** (kha) demonstrated by the deaf signers of Mary Chapman School for the Deaf, Yangon. These videos were converted into the corresponding image frames and these images were trained and classified by using 12 different transfer learning models. The results of epoch 20 and epoch 50 using these transfer learning models are compared and discussed in the Section VI.

Ni Htwe Aung and Su Su Maung are with the Department of Computer Engineering and Information Technology, Yangon Technological Technology, Insein, Yangon, 11011 Myanmar, Corresponding author e-mail: nhadec@gmail.com.

Ye Kyaw Thu is with LST Lab., NECTEC, Thailand, Corresponding author email: yktnlp@gmail.com

Swe Zin Moe and Hlaing Myat Nwe are with the Department of Information Science, University of Technology (Yatanarpon Cyber City), Pyin Oo Lwin, Mandalay, 05081 Myanmar

Manuscript received December 21, 2019; accepted March 6, 2020; revised March 16, 2020; published online April 30, 2020.

## II. Sign Language

Sign Language (SL) is a language that is mostly used as a form of non-verbal communication method by the hearing-impaired persons to communicate with their environment. SL is also a vision-based communication tool because it is only based on the power of vision. SL uses the action which contains the movements of body, hands, arms, lips, head and facial expressions instead of using sounds. Moreover, SL is not a universal language because different sign languages are used in different countries. Sign Language can differ from region to region, countries to countries and continents to continents. Moreover, each Sign Language has its own grammar structure and it is very different from the grammar structures of spoken languages.

Sign Language can be used for three different forms [2]. The first one is fingerspelling which is the sign used to describe each of the alphabet and number. It contains only hand movement and is mainly used to spell out the names of people, city, places, organizations, and for others which have no signs for these. The second one is word level sign which has the associated sign for each word of the vocabulary and it is used in combination with hand and facial expression. The third and essential one is non-manual sign which involves facial expressions, tongue, mouth, eyebrows, eyes, chin and body movement.

## III. Myanmar Sign Language (MSL)

### A. MSL Overview

Myanmar Sign Language (MSL) is the essential communication tool for the Myanmar deaf people. Same as other sign languages, MSL has different grammatical structure from Myanmar Language. As shown in Fig. 1, MSL is also implemented with manual and non-manual components like other SLs such as American Sign Language, British Sign Language, etc. The manual components which contain only hand shapes, hand position and hand movements, are mainly used to describe each letter of Myanmar and English alphabets, numbers and symbols. These manual signs (which is also called "Fingerspelling") are specially used in teaching the alphabets for the deaf children in primary education. To show feelings and meanings, non-manual components are used with facial expressions, movement of head, tension, eyebrows, eyelid, tongue, mouth and body [3]. As discussed in Section I, there are mainly two different sign languages in Myanmar: one is used in Northern part of Myanmar and the other one is used in Southern part of Myanmar. There are four deaf schools for the children in Myanmar [4]:

- Mary Chapman School for the Deaf, Yangon (est. 1904),
- School for the Deaf Children, Tamwe, Yangon (est. 2014),
- School for Deaf Children, Mandalay (est. 1964) and
- Immanuel School for the Deaf, Kalay (est. 2005).

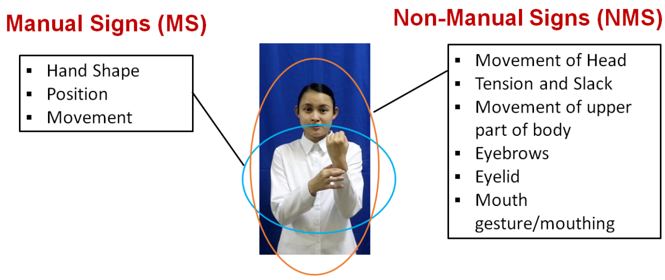
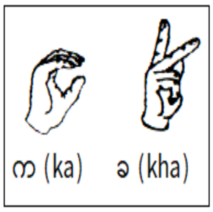


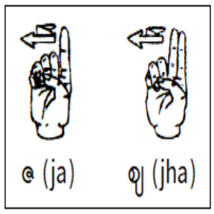

Fig. 1: Structure of Myanmar Sign Language

### B. Myanmar Fingerspelling

Myanmar deaf people use Myanmar fingerspelling which is the basic part of Myanmar Sign Language to represent Myanmar consonant, vowels and numbers and to spell out names of people, cities, places, organizations, and other words for which no sign exists in Myanmar sign language. It is also used in combination with existing signs to emphasis the concept or meaning. Myanmar fingerspelling characters contain static sign which represents a single image and dynamic sign which represents a sequence of multiple images. Only in 33 Myanmar consonants, there are 31 static signs and 2 dynamic signs. An example of static and dynamic Myanmar fingerspelling consonants is shown in Fig. 2 using Myanmar fingerspelling keyboard developed by Ye Kyaw Thu et al. [5] . Moreover, there are two different fingerspelling signs in Myanmar Sign Language; one is used in Mary Chapman School for the Deaf (Yangon) and another is used in School for the Deaf (Mandalay) and School for the Deaf (Tamwe, Yangon). The main difference is found in vowels, medial and symbol [6]. Only focuses on 33 Myanmar consonants fingerspelling characters, there are 12 different signs among these schools. An example of Myanmar consonants fingerspelling character difference between these schools is shown in Fig. 3.



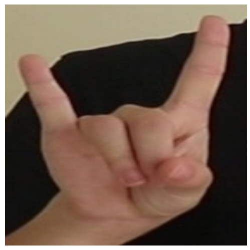

(a) Static Sign




(b) Dynamic Sign

Fig. 2: An example of static and dynamic Myanmar fingerspelling consonants

(a) Tamwe Sign for ယ (ya)



(b) Mary Chapman Sign for ယ (ya)

Fig. 3: An example of MSL fingerspelling character difference between School for the Deaf (Tamwe) and Mary Chapman School for the Deaf

## IV. Related Work

About 20 years ago, Sign Language recognition system was developed and its first publication had emerged in the beginning of the 90s. Most of the SL recognition approaches needed the use of expensive hardware devices such as gloves, 3D camera or low noise and high resolution images. The first Myanmar Fingerspelling Recognition System, which contains 30 static and opened finger images of Myanmar alphabets 'က' (ka) to 'အ' (a), was developed by Wah Wah et al. using Canny Edge detection and Artificial Neural Network (ANN). This system obtained the accuracy of 96% [7]. Thiri Min et al. also developed a video based MSL recognition system for 30 static and opened finger images of Myanmar alphabets 'က' (ka) to 'အ' (a) using Fast Hartley Transform (FHT) for feature extraction and Multilayer Perceptron (MLP) for classification and the system provided the classification accuracy of 96% [8]. In our first previous Myanmar Fingerspelling Recognition System, 31 static, opened and closed finger images were used and provided the higher accuracy of 97% using Random Forest Classifier [9]. The second previous approach of SL recognition for Myanmar Numbers used '၀' (0) to '၁၀' (10) images which represent the number signs used in Mary Chapman School. This approach tested and evaluated using the three different Support Vector Machine (SVM) Classifiers and provided the highest accuracy of 87% [10].

## V. Transfer Learning

Transfer learning is also a machine learning approach where the knowledge of the previous task can be used on the new related task. Transfer learning is different in building and training the model from traditional machine learning. Traditional machine learning is isolated and cannot consider past learned knowledge in other tasks and it can break down when there is no sufficient labeled data for the task of training a reliable model. In transfer learning, learning process can be faster, more accurate and less training data are needed and exist labeled data of some related task. Since 1995, transfer learning gets more attraction by researchers in different names such as learning learn, life-long learning, knowledge transfer, inductive transfer, multi-task learning, knowledge consolidation, context-sensitive learning, knowledge-based inductive bias, meta learning, and incremental/cumulative learning [11]. Among them, multi-task learning has closely related learning technique to transfer learning [12]. However, the roles of the source and target tasks in transfer learning are not symmetric as in multi-task learning [13] [14] [15].

We make a brief discussion for the well-known transfer learning models used in this experiment as follows:

1) VGG16: VGG is a deep convolutional network for object recognition created by Visual Geometry Group (VGG) which achieved the $1^{st}$ runner-up in ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014 [16]. VGG16 is a convolutional neural network which consists of 16 layers of deep neural network proposed by Karen Simonyan and Andrew Zisserman in 2015.Its architecture is simple because it is not using very much hyper parameters. It always uses the fixed size 224x224 RGB image as input and the image is passed through the stack of convolutional layers where 3x3 filters with stride of 1 in convolutional layer and uses the same padding in pooling layers 2x2 with stride of 2 [17].
2) VGG19: This network is also characterized using 3x3 convolutional layer stack and uses two fully-connected layers like VGG16. Unlike VGG16, VGG19 neural network consists of 19 layers of deep neural network. Although the size of VGG16 network with fully connected nodes is 533MB, the size of VGG19 network is 574MB. Moreover, VGG19 has more weight than VGG16 [17].
3) ResNet50: Residual Neural Network (ResNet) was the winner of ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2015 proposed by He et al. ResNet50 is also a convolutional neural network which consists of 50 layers of deep neural network. Even though it is much deeper than VGG16 and VGG19 models, its size is smaller due to the global average pooling rather than the fully-connected layers [18] [19].
4) InceptionV3: The Inception deep convolutional micro-architecture was first introduced as GoogLeNet by Szegedy et al. in 2014 and its goal is to work as a multi-level feature extractor by computing 1x1, 3x3 and 5x5 convolutions within the same network [20]. The subsequence appearance have been called Inception vN where N is the version number. Therefore, Inception V3 is the third version which includes the additional factorization ideas developed by Szegedy et al. in 2015 [21]. The weights for Inception V3 are smaller than both VGG and ResNet models and the network is 48 layers deep.
5) InceptionResNetV2: Inception-ResNet combines one Inception and Residual Networks and is able to give the higher performance and higher accuracies at a lower epoch. InceptionResNetV2 is a sub-version of Inception ResNet and it is introduced by Szegedy et al. in 2016 [22]. Its computational cost is similar to the Inception-v4 model and network is 164 layers in deep neural network.
6) Xception: Xception is an extension of Inception modules that have been replaced with depthwise separable convolutions. Xception has same parameter as Inception-v3 but it has the smallest weight serialization with size of 91MB [23].
7) MobileNet: It is a lightweight deep convolutional neural network that uses depthwise separable convolutions. Therefore, it can reduce the number of parameters significantly compared with other normal convolutional networks. Although, it is a smaller and faster network than the other, it needs very low maintenance [24].
8) DenseNet: DenseNet (Dense Convolutional Network), which connects each layer to every other layer in the feed forward fashion, was introduced by Cornell University, Tsinghua University and Facebook AI Research (FAIR) and got the best paper awards [25]. With dense connection, it achieves fewer parameters and high accuracy than the other models. Whereas traditional convolutional networks with "L" layers have "L" connections - one between each layer and its subsequent layer, DenseNet has $L(L+1)/2$ direct connections [26]. DenseNet architecture is highly efficient in parameter use and computation time [27].
9) NasNetMobile: NasNetMobile is generated based on a reinforcement learning technique, which known as AutoML (Automated Machine Learning) [28], and specifically designed to perform well on the popular Imagenet dataset [29]. This model achieves the satisfied results with smaller model size and lower complexity.

10) NasNetLarge: NasNetLarge is also a convolutional neural network which has learned rich feature representations for a wide range of images [29]. The network is trained on more than one million images from the ImageNet database. The input size of NasNetLarge model is 331x331 where NasNetMobile uses 224x224 input size.

## VI. Classification Approach

### A. Dataset and Preprocessing

Fig. 4: Static Myanmar Fingerspelling Consonants used in Mary Chapman School for the Deaf Children, Yangon

Firstly, the proposed system recorded MSL videos for Myanmar consonants demonstrated by 12 male and female deaf signers who have age between 12 and 27 from Mary Chapman School for the Deaf (Yangon) using Canon 200D camera with the resolution of 1,920x1,080. At this stage, all videos were taken in various indoor environments under normal lighting conditions using three different background colors; light-yellow, white and blue and two different color clothes, white and black, are used. All videos have recorded by capturing the upper part of the body of deaf signers and the recorded videos were converted into multiple image frames. Then the areas of hand region were segmented from these image frames manually. The segmented hand-only region are resized into 224x224 images and then passed the transfer learning model for training, validation and testing processes. The dataset of this proposed system contains 23,915 images of 31 static Myanmar fingerspelling consonants from 'က' (ka) to 'အ' (a) which are currently using in Mary Chapman School for the Deaf Children, Yangon. Myanmar fingerspelling images used in this proposed system are shown in Fig.4.

### B. Training and Validation

In the training stage, we used 19,151 images of Myanmar fingerspelling consonants (80% of total data) as the training dataset to train and fit the 12 different transfer learning models which are very popular in machine learning: VGG16, VGG19, ResNet50, InceptionV3, Xception, InceptionResNetV2, MobileNet, DenseNet121, DenseNet169, DenseNet201, NASNetLarge and NASNetMobile. For each model, the pretrained model was loaded and trained a new classifier on top for Myanmar consonant fingerspelling images and the architecture of the proposed system is shown in Fig.5.

To evaluate and ensure the performance skill of the transfer learning model, validation process is also used in this experiment. In validation stage, 2,382 images that are not contained in the training dataset are used as the validation data. In our experiment, we train and validate the transfer learning models using epoch 20 and epoch 50 and then compared the training and validation accuracy for each epoch and for each model. Fig. 6, Fig. 7, Fig. 8, Fig. 9 and Fig. 10 displays the training and validation accuracy graphically for some models (VGG16, ResNet50, MobileNet, DenseNet121, NASNetMobile) used in this experiment for epoch 20. Fig. 11, Fig. 12, Fig. 13 and Fig. 14, Fig. 15 also displays the training and validation accuracy for epoch 50 of these models.

### C. Recognition

In classification stage, 2,382 images that are not contained in training and validation datasets, are used for testing each model trained with epoch 20 and 50 and compared the prediction results for each model. The accuracy comparison graph for each model using two different epochs is shown in Fig.16. According to this figure, all of the models except InceptionV3, Xception and NASNetLarge have the higher accuracy. In this experiment, we made prediction for each image in the test dataset and calculate the percentage of true prediction of each model on new data. The prediction accuracy for each model of epoch 20 and epoch 50 is clearly shown in Table I. The highest accuracies among these models are highlighted in the table. According to this table, it can be seen that the best accuracy score of 94% was achieved in VGG16, ResNet50 and MobileNet models with epoch 50. DenseNet121 also achieved the higher accuracy of 92% on both epoch 20 and 50. VGG19, DenseNet169 and DenseNet201 also achieved the encouraging result for both epoch 20 and epoch 50. It can be clearly seen that the accuracy is improved significantly in ResNet50, InceptionResNetV2 and NASNetMobile using epoch 50. InceptionV3 and Xception models have the comparative

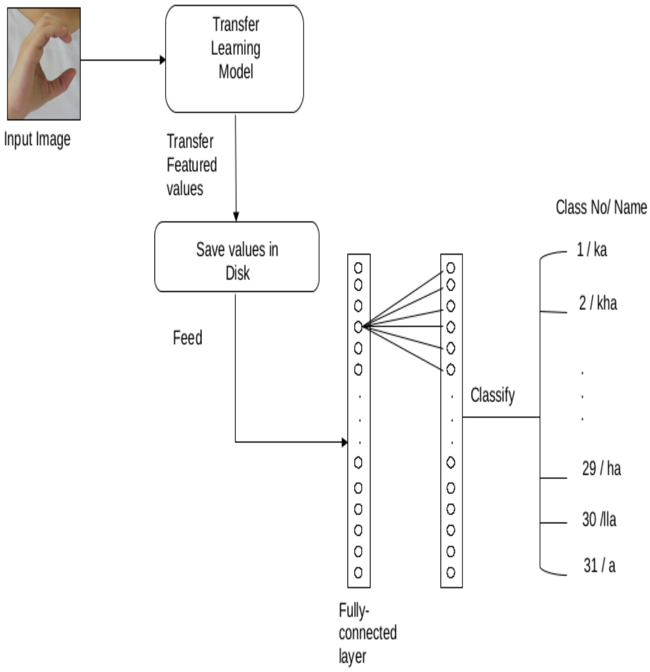



Fig. 5: Architecture of the proposed system

results for both epoch 20 and epoch 50 but the accuracy of NASNetLarge model remains unchanged for both epoch 20 and 50.

TABLE I: Accuracy (%) for each Transfer Learning Model

| No. | Model | Epoch20 | Epoch50 |
|---|---|---|---|
| 1 | VGG16 | 81% | **94**% |
| 2 | VGG19 | 83% | 89% |
| 3 | ResNet50 | 59% | **94**% |
| 4 | InceptionV3 | 37% | 44% |
| 5 | Xception | 50% | 40% |
| 6 | InceptionResNetV2 | 57% | 73% |
| 7 | MobileNet | 90% | **94**% |
| 8 | DenseNet121 | 92% | 92% |
| 9 | DenseNet169 | 84% | 83% |
| 10 | DenseNet201 | 82% | 85% |
| 11 | NASNetMobile | 49% | 67% |
| 12 | NASNetLarge | 42% | 30% |

## VII. Conclusion

Although there are some limitations in the datasets of our experiment, we obtained the encouraging result with very few preprocessing stages for different background colors, different clothes color, different lighting condition and different hand locations. Moreover, our system is capable of classifying 31 Myanmar fingerspelling consonants for both opened and closed fingers without the need for any special expensive hardware devices such as gloves, 3D cameras or sensors. In the near future, we intended to develop a Myanmar Fingerspelling recognition system for all Myanmar fingerspelling consonants, vowels and symbols including both static and dynamic signs by applying deep neural network.

## Acknowledgment

We would like to give special thanks to the principals, teachers, MSL translators and students from Mary Chapman School for the Deaf (Yangon), School for the Deaf (Tamwe, Yangon) and School for the Deaf (Mandalay). We would like to thank JICA EEHE Project (Project for Enhancement of Engineering Higher Education in Myanmar) for their supporting of research fund to our research. We would also like to thank all participants for their kind contributions to our research. We give special thanks to Google Inc. for publicly available some Transfer Learning models and Tensorflow library.

## References

[1] Disability population in Myanmar, 2014.
[2] Ilan Steinberg, Tomer M. London, Dotan Di Castro, “Hand Gesture Recognition in Images and Video”, Technion-Israel Institute of Technology, 2003
[3] Y. Y. Swe, Myanmar Sign Language Basic Conversation Book, 1st Edition ed., Department of Social Welfare, Ministry of Social Welfare, Relief and Resettlement, Department of Social Welfare, Japan International Cooperation Agency, August 2009.
[4] ”Burmese Sign Language (or) Myanmar Sign Language”. Available: https://en.wikipedia.org/wiki/Burmesesignlanguage.
[5] Ye Kyaw Thu, S. A. W. M. and URANO, Y. “Direct Keyboard Mapping (DKM) Layout for Myanmar Fingerspelling Text Input (study with Developed Fingerspelling Font)”
[6] ”Text book of Speaking and Myanmar Sign Communication”, Mary Chapman School for the Deaf, 1988.
[7] W. Wah, Myanmar Sign Language Recognition System Using Artificial Neural Network, December, 2014.
[8] Thiri Min, Thanda Aung, ”Video Based Myanmar Sign Language Recognition System,” in 12th National Conference on Science and Engineering, Mandalay, Myanmar, 2019.
[9] Ni Htwe Aung, Ye Kyaw Thu, Su Su Maung, ”Feature Based Myanmar Fingerspelling Image Classification Using SIFT, SURF and BRIEF,” in In Proceedings of the 17th International Conference on Computer Applications (ICCA), Yangon, February, 2019.
[10] Ni Htwe Aung, Su Su Maung, Ye Kyaw Thu, ”Sign Language Recognition for Myanmar Number Using Three Different SVM Classifiers,” in 12th National Conference on Science and Engineering, Yangon, June, 2019.
[11] S. Thrun and L. Pratt, Learning to learn, L. P. Sebastian Thrun, Ed., USA: Kluwer Academic Publishers,, 1998.
[12] R. Caruana, ”Multitask learning,” in Machine Learning, vol. 28(1), Netherlands, KluwerAcademicPublishers, 1997, pp. 41-75.
[13] Ying Lu, Transfer Learning for Image Classification, Université de Lyon, 2017, p. 71.
[14] Abhishek Kumar and Hal Daume, ”Learning Task Grouping and Overlap in Multi-task Learning,” in In Proc. 29th International Conference on Machine Learning (ICML), NY, USA, 2012.
[15] Yu Kong, Ming Shao, Kang Li and Yun Fu, ”Probabilistic LowRank Multitask Learning,” in IEEE Transactions on Neural Networks Learning Systems, 2017.
[16] ”ImageNet Large Scale Visual Recognition Challenge 2014 (ILSVRC2014),” [Online]. Available: http://www.imagenet.org/challenges/LSVRC/2014/results.
[17] Karen Simonyan and Andrew Zisserman, Visual Geometry Group, ”Very Deep Convolutional Networks For Large-Scale Image Recognition,” in International Conference on Learning Representations, 2015.
[18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, ”Deep Residual Learning for Image Recognition,” 2015.
[19] Kaiming He, Xiangyu Zhang,Shaoqing Ren, Jian Sun, ”Deep Residual Learning for Image Recognition,” in In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
[20] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed,Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, ”Going deeper with convolutions,” in eprint arXiv 1409.4842, 2014.

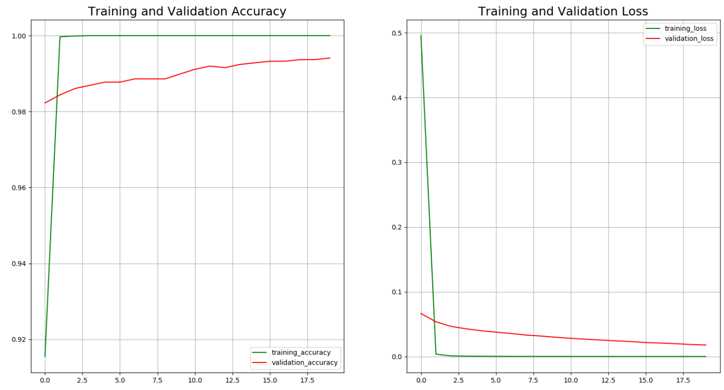



Fig. 6: Training and Validation Accuracy of epoch 20 for VGG16

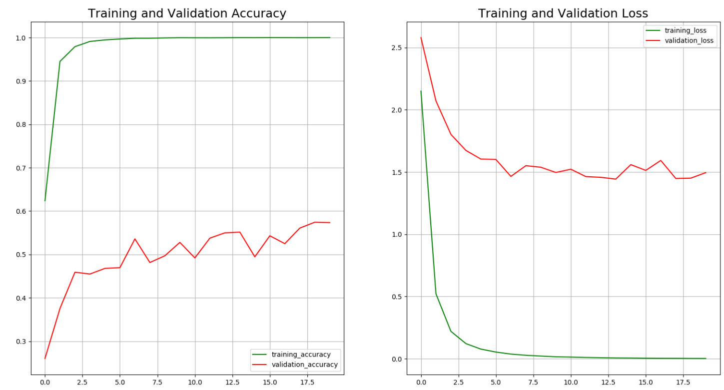



Fig. 7: Training and Validation Accuracy of epoch 20 for ResNet50

[21] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, "Rethinking the Inception Architecture for Computer Vision," in eprint arXiv:1512.00567, 2015.
[22] Szegedy, Christian, Sergey Ioffe, Vincent Vanhoucke, and Alexander A. Alemi., "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," in AAAI, 2017.
[23] F. . Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in eprint arXiv:1610.02357, 2016.
[24] Howard, Andrew G.; Zhu, Menglong; Chen, Bo; Kalenichenko, Dmitry; Wang, Weijun; Weyand, Tobias; Andreetto, Marco; Adam, Hartwig, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," in eprint arXiv:1704.04861, 2017.
[25] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei Fei, "ImageNet Large Scale Visual Recognition Challenge," in International Journal of Computer Vision (IJCV), 2015.
[26] Huang, Gao; Liu, Zhuang; van der Maaten, Laurens; Weinberger, Kilian Q., "Densely Connected Convolutional Networks," in eprint arXiv 1608.06993, 2016.
[27] Pleiss, Geoff; Chen, Danlu; Huang, Gao; Li, Tongcheng; van der Maaten, Laurens; Weinberger, Kilian Q., "Memory-Efficient Implementation of DenseNets" in eprint arXiv:1707.06990, 2017.
[28] Zoph, Barret; Le, Quoc V., "Neural Architecture Search with Reinforcement Learning," in eprint arXiv:1611.01578, 2016.
[29] Zoph, Barret, Vijay Vasudevan, Jonathon Shlens, and Quoc V. Le, "Learning Transferable Architectures for Scalable Image Recognition," in eprint arXiv:1707.07012 , 2017.

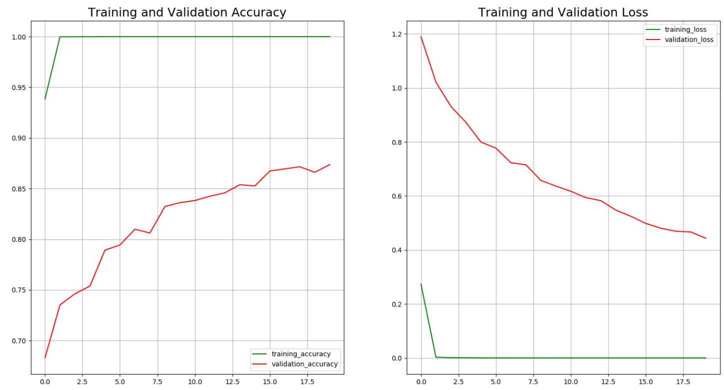



Fig. 8: Training and Validation Accuracy of epoch 20 for MobileNet

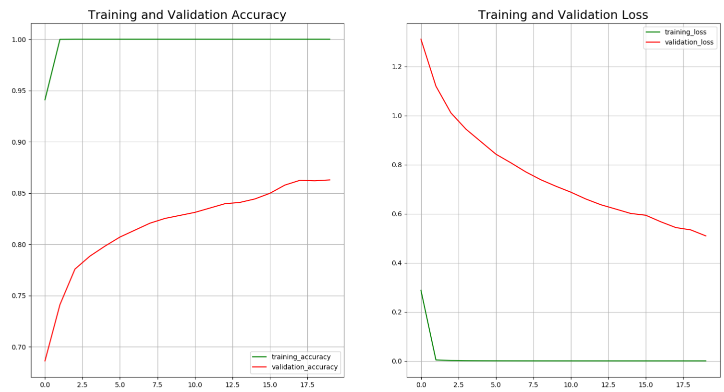



Fig. 9: Training and Validation Accuracy of epoch 20 for DenseNet

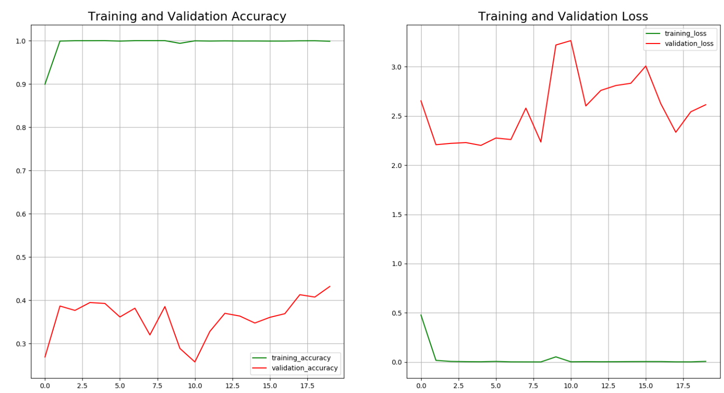



Fig. 10: Training and Validation Accuracy of epoch 20 for NASNetMobile

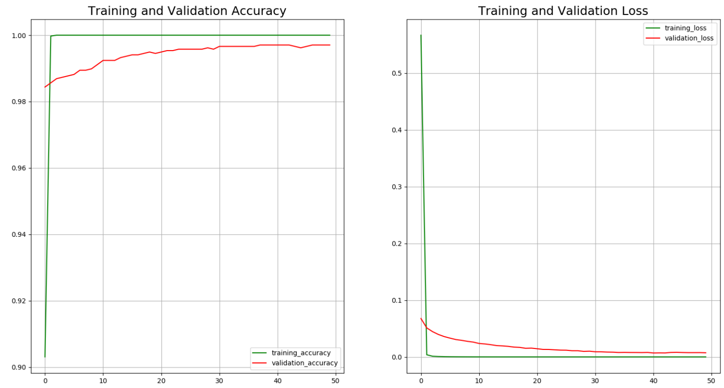



Fig. 11: Training and Validation Accuracy of epoch 50 for VGG16

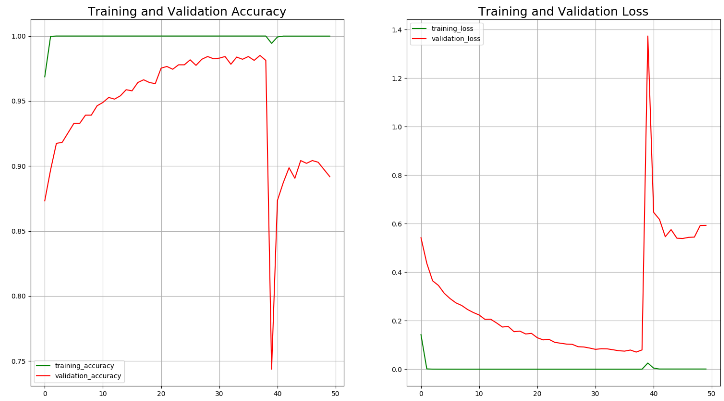



Fig. 12: Training and Validation Accuracy of epoch 50 for ResNet50

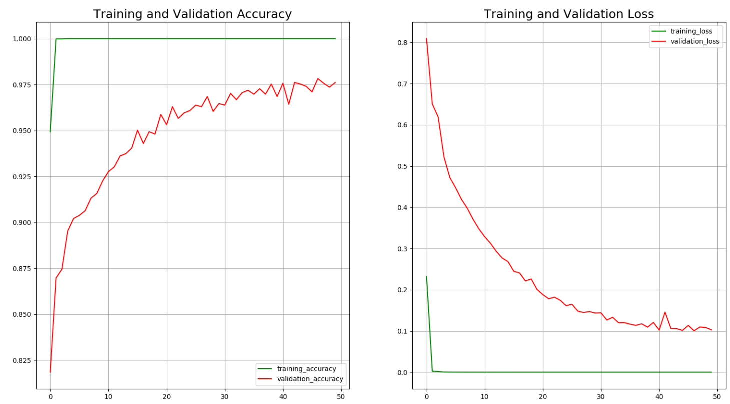



Fig. 13: Training and Validation Accuracy of epoch 50 for MobileNet

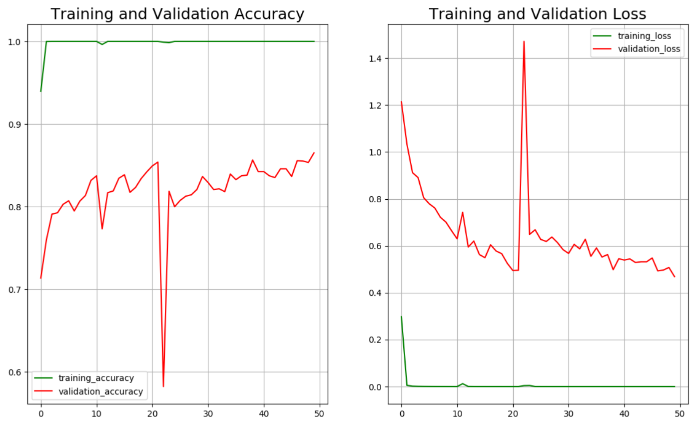



Fig. 14: Training and Validation Accuracy of epoch 50 for DenseNet

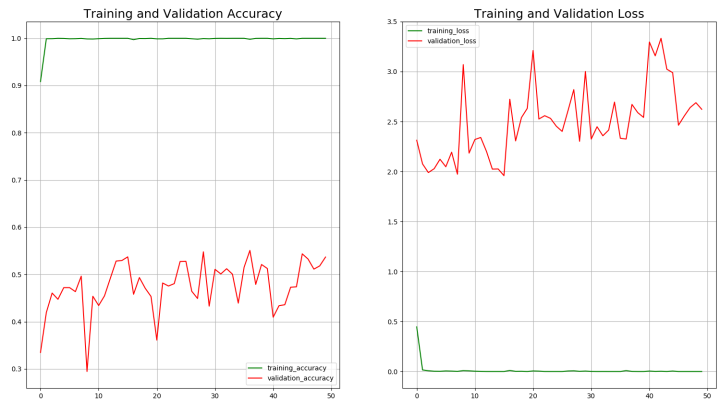



Fig. 15: Training and Validation Accuracy of epoch 50 for NasNetMobile

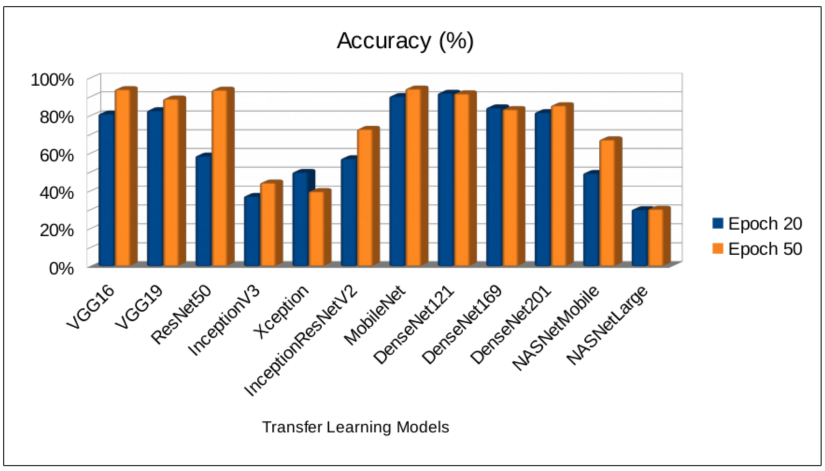



Fig. 16: Classification accuracy of 12 Transfer Learning models (two trainings; 20-epoch training and 50-epoch training)

**Ni Htwe Aung** received the B.E. degree in Information Technology at Technological University (Monywa) in 2007, and M.E degrees in Information Technology from Technological University (Monywa) in 2010. In 2011, she joined the Department of Information Technology, Technological University (Monywa), as an Instructor, and in 2018 she became an Assistant Lecturer at the Department of Computer Engineering and Information Technology, Yangon Technological University. She is also a Ph.D candidate in Computer Engineering and Information Technology (CEIT) Department of Yangon Technological University (YTU). She focused her efforts on her doctoral thesis research of Real-time Myanmar Fingerspelling Recognition.

**Ye Kyaw Thu** is a Visiting Professor of Language & Semantic Technology Research Team (LST), Artificial Intelligence Research Unit (AINRU), National Electronic & Computer Technology Center (NECTEC), Thailand and Head of NLP Research Lab., University of Technology Yatanarpon Cyber City (UTYCC), Pyin Oo Lwin, Myanmar. He is also a founder of Language Understanding Lab., Myanmar and a Visiting Researcher of Language and Speech Science Research Lab., Waseda University, Japan. He is actively co-supervising/supervising undergrad, masters' and doctoral students of several universities including MTU, UCSM, UCSY, UTYCC and YTU.

**Su Su Maung** received the B.E. degree in Information Technology from Mandalay Technological University (MTU), in 2002, and the M.E degrees in Information Technology from Yangon Technological University (YTU), in 2004. She obtained her Master's Degree in Information Engineering from the Faculty of Engineering, KMIL, Bangkok, Thailand (2006), and her Doctoral Degree also in Information Technology from Mandalay Technological University (MTU), in 2008. She worked as an Associate Professor at University of Technology (Yadanarpon Cyber City) (2010-2014). She is now working as an Associate Professor in Computer Engineering and Information Technology (CEIT) Department of Yangon Technological University (YTU) since 2015. Her research interest include image processing, signal processing, machine learning.

**Swe Zin Moe** is currently an Assistant Lecturer at Myanmar Institute of Information Technology (MIIT), Mandalay, Myanmar and also a Ph.D candidate at University of Technology (Yatanarpon Cyber City), Pyin Oo Lwin, Myanmar. Her current doctoral thesis research focuses on machine translation between Myanmar sign language and Myanmar written text. She is interested in the general and related problems of natural language processing (NLP) such as machine translation, big data analysis and deep learning.

**Hlaing Myat Nwe** is a PhD candidate of University of Technology (Yatanarpon Cyber City), Pyin Oo Lwin, Myanmar. A native of Myanmar, she holds a master degree of Information Science and Technology, and a bachelor degree of Information Science and Technology from University of Technology (Yatanarpon Cyber City), Myanmar. Her research interests include human-computer interaction, natural language processing and Artificial Intelligent. She has been working to find efficient and user-friendly text input interfaces for Myanmar Sign Language.