CTU Journal of Innovation and Sustainable Development
ISSN 2615-9422 | e-ISSN 2815-5602

DOI:10.22144/ctujoisd.2023.042

Analysis of printed document identification based on Deep Learning

Nguyen Dinh Thong, Phu Quang Nguyen*, and Mai Hoang Bao An*

School of Computer Science and Engineering, International University, Viet Nam National University Ho Chi Minh City, Viet Nam

**Corresponding author (nqphu@hcmiu.edu.vn, mhban@hcmiu.edu.vn)*

Article info.

Received 16 Jul 2023
Revised 18 Sep 2023
Accepted 3 Oct 2023

Keywords

GAN, Mixup Augmentation, Mobile App, Printed Source Identification, ResNet

ABSTRACT

In this study, we investigate the effectiveness of ResNet, a deep neural network architecture, for a deep learning approach to address the problem of printed document identification. ResNet is known for its ability to handle the vanishing gradient problem and learn highly representative features. Multiple variations of ResNet have been applied, including ResNet50, ResNet101, and ResNet152, which provide the backbone architecture of our classification model and are trained on a comprehensive dataset of microscopic printed images containing some microscopic printing patterns from various source printers. We also incorporate Mix-up augmentation, a technique that generates virtual training samples by interpolating pairs of images and labels, to further enhance the performance and generalization capability of the model. The experimental results showed that ResNet101 and ResNet152 variants outperformed in accurately distinguishing printer sources based on microscopic printed patterns. We developed a mobile app to test the feasibility of our findings in practice. In conclusion, this study aims to lay the groundwork for creating a sufficiently pre-trained model with accurate performance of identification that can be deployed on mobile devices to detect the printed sources of documents.

1. INTRODUCTION

To fight against the counterfeit industry effectively, innovative solutions are needed that can accurately authenticate and classify genuine products. This research focuses on developing a classification framework that leverages advanced deep learning techniques, to identify and categorize authentic products based on the use of printed graphical codes stuck on each product (performing authentication for printed documents). By enhancing our ability to distinguish genuine items from counterfeits, we can protect consumer trust, safeguard industries, and mitigate the detrimental effects of counterfeit trade on the global economy and public health.

This paper proposes a comprehensive approach that combines microscopic printing analysis and deep learning techniques to classify the source of printed documents, responsible for producing microscopic printing patterns carved on various products. By accurately identifying the source of printed patterns, we establish a reliable link between the printing pattern and the genuine manufacturer, facilitating the authentication process and aiding in the fight against counterfeiting.

The subsequent chapters will delve into the existing literature, methodologies, experimental results, and analysis, leading to the conclusion and potential future directions for research in this vital area of classification for authenticating products.

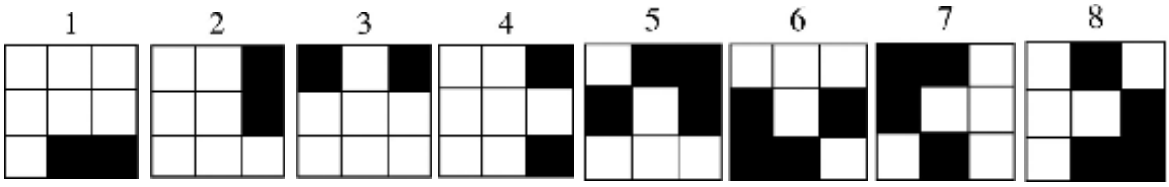


Figure 1. Microscopic printing patterns applied

2. MATERIALS AND METHOD

In this section, we describe the method used in this paper for the identification of printed documents, which uses various ResNet (He et al., 2016) based models as the backbone. To enhance the training performance, we also incorporate Mixup augmentation to generate more similar data, and then perform the training for different patterns using a combination of different paper and printers.

2.1. Dataset

In this study, we examined the same dataset with authors from (Nguyen et al., 2021), (Nguyen et al., 2021), and (Vo et al., 2022) which contain microscopic images of eight distinguished patterns printed using three different printing technologies on two types of papers. This dataset was achieved with numerous hours of labor following these steps:

First, two commonly used printing methods were employed, Offset and Xerography, to print on two different surfaces: (1) uncoated natural paper and (2) paper coated with one or more layers. Offset printing, a lithographic technique, involves transferring the inked image from a printing plate (containing the desired pattern) to a rubber blanket and then onto the printing surface through a process called "Offset" (Kipphan, 2001). This work used two specific offset printing techniques: (1) Conventional Offset Printing, where water and additives dampen the image-carrying plate, and (2) Waterless Offset Printing, which employs a silicone layer that repels ink, covering the non-inked areas of the printing plate. Whereas, Xerography, also known as laser printing, is a dry photocopying method. It uses a laser beam to electrostatically charge the photoconductor, attracting inks, powders, or liquid toners that carry an opposite charge to the photoconductor surface. Subsequently, the transferred ink is fixed to the substrate by subjecting it to elevated temperatures (Kipphan, 2001).

Next, eight distinct patterns outlined in Figure 1 were produced using the previously outlined setups. These patterns were chosen based on an evaluation of microscopic observations. The sample collection process involved using an optical Zeiss Microscope with an AxioCam camera. The outcomes following this step can be seen in Figure 2.

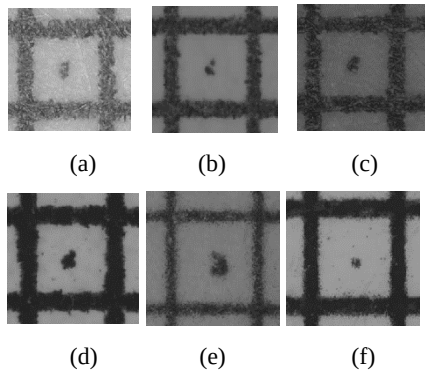


Figure 2. The uncoated (a, c) and coated (b, d) papers with conventional (a, b) and waterless (c, d) offset printing technologies, uncoated (e) and coated (f) papers with xerography printing technology

To create a diverse and representative dataset, printer source images were gathered from various sources, including different manufacturers, printing technologies, and printing papers. The dataset collection process involved obtaining eight patterns for each printer model and substrate combination. For consistency and reliability, one hundred copies were produced for each pattern of each printer type. These copies were then captured using a Zeiss Microscope to ensure high-resolution images for further analysis.

Table 1. Details of the Experimented Dataset

Printing Technology	Printing Paper	Number of samples per pattern
Conventional Offset	Uncoated	100
Conventional Offset	Coated	100
Waterless Offset	Uncoated	100
Waterless Offset	Coated	100
Xerography	Uncoated	100
Xerography	Coated	100

As shown in Table 1. Details of the Experimented Dataset, the dataset comprises 4800 samples, with

each sample corresponding to a specific combination of design, printer type, and paper type. This comprehensive dataset allows for an in-depth investigation of the patterns and their variations across different printers and substrate.

2.2. Data Preprocessing

Before proceeding with the training, several preprocessing steps were applied to the dataset. These steps aimed to enrich the source of images, remove the redundant noise, and hence improve the view of the learning model for subsequent classification tasks. The preprocessing pipeline comprised of the following steps:

GAN-based Augmentation: To augment the dataset and increase its diversity, we used Generative Adversarial Networks (GANs) (Isola et al., 2017). GANs are deep learning models that consist of a generator and a discriminator. The generator learns to generate realistic synthetic data samples, while the discriminator learns to distinguish between real and synthetic samples. Figure 3 describes the GAN model implementation on our dataset so that we could generate additional synthetic images that closely resemble the original printer source images. This augmentation technique helps to increase the variability of the dataset and hence improve the power of learning models.

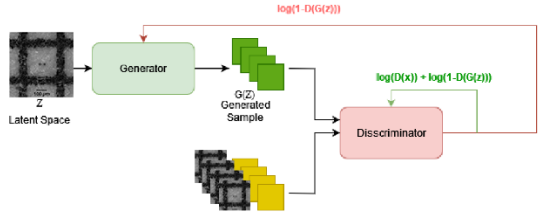


Figure 3. GAN architecture for generating more training images

Augmentation: Mix-up (Zhang et al., 2017) augmentation is another technique that we use to create new augmented samples. It involves taking a weighted average of the input images and their corresponding labels. By blending samples together, mix-up augmentation introduces smooth transitions between different classes and helps to improve the model's robustness and generalization. We applied mix-up augmentation to the dataset (Figure 4), creating augmented samples that combine the characteristics of multiple printer sources.

These preprocessing techniques, including GAN-based augmentation, and mix-up augmentation, helped enhance the dataset's diversity, improve the model's ability to generalize, and isolate the relevant microscopic printing patterns for accurate classification. The preprocessed dataset was then used for training the deep learning models with a list of selected ResNet variants as the backbone architecture.

Figure 4. Mix-up example

2.3. Model Architecture

The backbone architecture used in this study is based on the ResNet (Residual Network) models. ResNet is a deep learning architecture that has showed excellent performance in image classification tasks. It addresses the problem of vanishing gradients by introducing residual connections, which enable the network to effectively learn from deeper layers. In this study, we used various ResNet variants as the backbone architecture for our deep learning models. The following ResNet models were employed:

– **ResNet50**: This variant comprises 50 layers and has shown strong performance in image classification tasks. It includes residual blocks with skip connections, enabling the network to learn deep representations effectively.

– **ResNet101**: Similar to ResNet50, ResNet101 has 101 layers. It provides a deeper network architecture, allowing for more complex feature extraction and representation.

– **ResNet152**: ResNet152 is an even deeper variant with 152 layers. It offers increased model capacity and the ability to capture more intricate patterns and details in the input images.

– **ResNeXt50_32x4d**: ResNeXt (Xie et al., 2017) models introduce a cardinality parameter, which allows for more extensive network connectivity. ResNeXt50_32x4d has 50 layers and a cardinality of 32, enhancing its feature learning capabilities.

– **ResNeXt101_32x8d**: this variant provides a deeper and more connected architecture compared to ResNeXt50_32x4d with 101 layers and a cardinality of 32. It can capture more fine-grained features and exhibit improved classification performance.

– **ResNeXt101_64x4d**: this variant is similar to ResNeXt101_32x8d, but exhibits a higher cardinality of 64. This variant offers enhanced feature learning and representation abilities, making it suitable for complex classification tasks.

– **WideResNet50_2**: WideResNet (Zagoruyko et al., 2016) models increase the width (number of channels) of the network, providing a broader capacity for feature learning. WideResNet50_2 has double the number of channels compared to ResNet50, allowing for richer feature representations.

– **WideResNet101_2**: WideResNet101_2 has increased width and capacity with 101 layers. It can capture more diverse and detailed features from the input images.

These ResNet variants illustrated in Table 2 offer a range of architectural depths and complexities, allowing for flexibility in capturing and learning different levels of patterns and details in the printed source images. By utilizing these models as the backbone architecture, our deep learning models can effectively extract and leverage hierarchical features for the accurate classification of printed sources of documents based on their microscopic printing patterns.

Table 2. Architectures of Experimented ResNet Variants

Layer Name	conv1	conv2_x		conv4_x	conv5_x		
Output Size	112x112	56x56		14x14	14x14	1x1	
ResNet50	7x7, 64, Stride 2	3x3 Max pool, stride 2	$\begin{bmatrix}1\times1,64\\3\times3,64\\1\times1,256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1,128\\3\times3,128\\1\times1,512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,1024\end{bmatrix}\times6$	$\begin{bmatrix}1\times1,512\\3\times3,512\\1\times1,2048\end{bmatrix}\times3$	average pool, 1000-d fc, softmax
ResNet101			$\begin{bmatrix}1\times1,64\\3\times3,64\\1\times1,256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1,128\\3\times3,128\\1\times1,512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,1024\end{bmatrix}\times23$	$\begin{bmatrix}1\times1,512\\3\times3,512\\1\times1,2048\end{bmatrix}\times3$	
ResNet152			$\begin{bmatrix}1\times1,64\\3\times3,64\\1\times1,256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1,128\\3\times3,128\\1\times1,512\end{bmatrix}\times8$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,1024\end{bmatrix}\times36$	$\begin{bmatrix}1\times1,512\\3\times3,512\\1\times1,2048\end{bmatrix}\times3$	
ResNeXt50_32x4d			$\begin{bmatrix}1\times1,128\\3\times3,128,C=32\\1\times1,256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1,128\\3\times3,128,C=32\\1\times1,256\end{bmatrix}\times4$	$\begin{bmatrix}1\times1,512\\3\times3,512,C=32\\1\times1,1024\end{bmatrix}\times6$	$\begin{bmatrix}1\times1,1024\\3\times3,1024,C=32\\1\times1,2048\end{bmatrix}\times3$	
ResNeXt101_32x8d			$\begin{bmatrix}1\times1,128\\3\times3,256,C=32\\1\times1,256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1,128\\3\times3,256,C=32\\1\times1,256\end{bmatrix}\times4$	$\begin{bmatrix}1\times1,512\\3\times3,1024,C=32\\1\times1,1024\end{bmatrix}\times23$	$\begin{bmatrix}1\times1,1024\\3\times3,2048,C=32\\1\times1,2048\end{bmatrix}\times3$	
ResNeXt101_64x4d			$\begin{bmatrix}1\times1,64\\3\times3,64,C=64\\1\times1,256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1,128\\3\times3,128,C=64\\1\times1,256\end{bmatrix}\times4$	$\begin{bmatrix}1\times1,256\\3\times3,256,C=64\\1\times1,1024\end{bmatrix}\times23$	$\begin{bmatrix}1\times1,512\\3\times3,512,C=64\\1\times1,2048\end{bmatrix}\times3$	
WideResNet50_2			$\begin{bmatrix}1\times1,128\\3\times3,128\\1\times1,256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1,512\\3\times3,512\\1\times1,1024\end{bmatrix}\times6$	$\begin{bmatrix}1\times1,1024\\3\times3,1024\\1\times1,2048\end{bmatrix}\times3$	
WideResNet101_2			$\begin{bmatrix}1\times1,128\\3\times3,128\\1\times1,512\end{bmatrix}\times3$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,1024\end{bmatrix}\times4$	$\begin{bmatrix}1\times1,512\\3\times3,512\\1\times1,2048\end{bmatrix}\times23$	$\begin{bmatrix}1\times1,1024\\3\times3,1024\\1\times1,4096\end{bmatrix}\times3$	

2.4. Train-Validation-Test Split

After the preprocessing steps, we acquired 9600 images (4800 original + 4800 Generative images) and divided the dataset into three subsets: train, validation, and test. We used a ratio of 7:2:1 for these splits, respectively. The train set, comprising 70% of the dataset, was used to train our models and optimize their parameters. The validation set, comprising 20% of the dataset, served as a separate subset to monitor the model's performance during training and fine-tune hyperparameters. Finally, the test set, accounting for 10% of the dataset, was used to evaluate the models' generalization and assess their performance on unseen data.

2.5. Training procedure and evaluation metrics

For our experiments, we applied the aforementioned architectures for the Feature Extractor part of the framework. In the data loading part, the images are loaded with a batch size of 32. The model was trained for 100 epochs using SGD optimizer (Ruder

et al., 2016) with an initial learning rate 1×10^{-3} and decays by 0.1 after 15, 30, 45, and 60 epochs. We resized the image to 224x224 in both the training and testing process. Our study was built on Pytorch (Paszke et al., 2017) version 1.9.1 and conducted on a machine with NVIDIA RTX 3070Ti GPU. For evaluation, we used the macro CE_Loss-Score (Paszke et al., 2017) for measuring performances. We also recorded the number of parameters required from each model and the GPU RAM requirement with the same configuration.results and discussion

Additionally, it is important to consider the applications of these findings. Thus, we developed a mobile app shown in Figure 5 to test the feasibility of our study in practice. In the app, the user can upload an image of a printed pattern (Figure 5a), choose a prediction model (Figure 5b), and receive the predicted result (Figure 5c).

Table 3 represents the results of our experiment involving multiple ResNet variants on eight aforementioned microscopic printing patterns. The table includes accuracy and cross-entropy (CE) loss values for each pattern and provides valuable insights into the performance of the tested materials, printing technologies, and patterns. From these results, we perform a detailed analysis of the experiment.

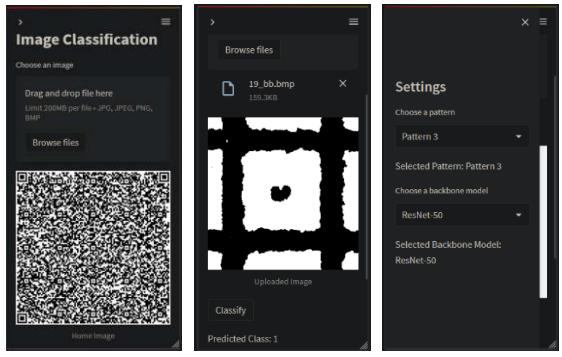
In this experiment, we focus on microscopic printing patterns and their imaging from several printing technologies on distinct substrates to evaluate the effectiveness of these patterns in terms of accuracy and CE loss. Eight distinct microscopic printing patterns, represented by Pattern 1 to Pattern 8, have diverse characteristics of structural variations. The materials used for printing will not be discussed here. The accuracy values measure the success rate of correctly identifying or matching the printed patterns. In comparing the results across the various models utilized in this research, we observe that the accuracy achieved is relatively high. The percentages provided range from 95.56% to 100% in comparison with the range from 66% to 80% in our previous study (Vo et al., 2022). This shows the effectiveness of the methodologies applied and the feasibility of using the deep learning approach for microscopic printer identification.

From the indicated values in the result table, we can observe that pattern 5 consistently shows high accuracy, ranging from 98.33% to 100%, indicating its robustness across different materials. Patterns 6 and 8 also exhibit high accuracy, consistently achieving 99.44% and 100% accuracy, respectively. Patterns 3 and 4 show relatively consistent accuracy values, ranging from 95.56% to 97.78% and 97.22% to 99.44%, respectively. Whereas, patterns 1, 2, and 7 display varied results, showing potential sensitivity to the materials or printing technologies.

Interestingly, the models that were trained on more complex patterns (pattern 5, 6, 7, and 8) outperformed their counterparts that were trained on simpler patterns (pattern 1, 2, 3, and 4). These models achieved better results in terms of lower Cross-Entropy (CE) loss, suggesting that increasing the complexity of the patterns contributes to improving the model's ability to classify the source printer accurately.

It was initially hypothesized that adding complexity might raise a challenge to the model, potentially leading to overfitting. However, the findings suggest the opposite: the increased complexity seems to have provided the models with more distinctive features to learn from, thereby enhancing their discriminative power.

Given these encouraging results, it can be concluded that this dataset and the proposed approach offer significant potential for real-world applications. The high accuracy and improved performance with higher pattern complexity suggest that even more complex patterns might lead to further improvements, marking a promising direction for future research.



(a) (b) (c)

Figure 5. Upload page (a), Choose model page (b), and Result page (c)

These findings make a substantial contribution to the field of source printer identification, particularly when based on microscopic printed documents. The results show the potential of deep learning techniques in this domain and pave the way for more sophisticated approaches in the future.

Additionally, it is important to consider the applications of these findings. Thus, we developed a mobile app shown in Figure 5 to test the feasibility of our study in practice. In the app, the user can upload an image of a printed pattern (Figure 5a), choose a prediction model (Figure 5b), and get the predicted result (Figure 5c).

Table 3. Accuracy and CE Loss for Different Patterns and Models on test set at best epoch

		ResNet 50	ResNet 101	ResNet 152	ResNeXt 50_32x4d	ResNeXt 101_32x8d	ResNeXt 101_64x4d	WideResNet 50_2	WideResNet 101_2
Pattern 1	Accuracy	**98.00%**	**98.00%**	**98.00%**	96.00%	96.00%	96.67%	96.00%	97.33%
	CE Loss	0.097	0.1211	**0.0777**	0.1315	0.1351	0.1509	0.1351	0.1218
Pattern 2	Accuracy	95.56%	96.11%	95.56%	95.56%	95.56%	**96.67%**	**96.67%**	96.11%
	CE Loss	0.1602	0.1461	0.135	0.1343	0.1579	0.1353	**0.1323**	0.1455
Pattern 3	Accuracy	96.67%	96.11%	95.56%	96.11%	95.56%	94.44%	**97.78%**	95.56%
	CE Loss	**0.1107**	0.1296	0.1375	0.1652	0.1487	0.1741	0.1272	0.1764
Pattern 4	Accuracy	97.78%	**99.44%**	98.89%	97.78%	97.22%	98.33%	96.11%	97.78%
	CE Loss	0.0769	6.5	**0.0653**	0.1039	0.1233	0.1012	0.1207	0.12
Pattern 5	Accuracy	98.33%	98.33%	98.89%	98.89%	98.89%	97.22%	99.44%	**100.00%**
	CE Loss	0.0915	0.0605	0.04	0.0495	0.1013	0.1088	0.0504	**0.037**
Pattern 6	Accuracy	**100.00%**	98.33%	99.44%	99.44%	99.44%	99.44%	99.44%	99.44%
	CE Loss	0.0171	0.0336	0.0211	0.024	0.0459	0.028	0.0261	**0.015**
Pattern 7	Accuracy	99.33%	99.33%	**100.00%**	99.33%	99.33%	98.00%	98.67%	99.33%
	CE Loss	0.0234	0.0487	**0.0104**	0.0207	0.0626	0.0699	0.0364	0.0359
Pattern 8	Accuracy	**100.00%**	98.00%	98.67%	**100.00%**	98.67%	98.67%	98.67%	99.33%
	CE Loss	0.0463	0.072	0.0564	**0.0256**	0.0718	0.0826	0.0474	0.0408

3. CONCLUSION

In conclusion, the research shows the effectiveness of deep learning approaches for microscopic printed document identification. The high accuracy achieved, and the superior performance of models trained on complex patterns highlight their potential in real-world applications. Contrary to initial expectations, more complicated patterns improved the potency to classify printers, offering distinctive features for better discrimination. We also developed a mobile app at the staging level to prove the feasibility of this approach. Future research should focus on developing new, secure patterns and implementing data adaptation techniques to handle actual data with noise. These findings contribute significantly to the field of printed document identification for authentication and lay the groundwork for enhanced anti-counterfeiting measures.

REFERENCES

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., ... & Lerer, A. (2017). Automatic differentiation in pytorch. *The 31st Conference on Neural Information Processing Systems* (NIPS 2017), Long Beach, CA, USA.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.

Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1125–1134.

Kipphan, H. (2001). *Handbook of print media: Technologies and production methods*. Springer Science and Business Media.

Nguyen, Q. P., Dang, N. T., Mai, A., & Nguyen, V. S. (2021). Features selection in microscopic printing analysis for source printer identification with machine learning. In *International Conference on Future Data and Security Engineering* (pp. 210–223). Springer.

Nguyen, Q.-T., Mai, A., Chagas, L., & Reverdy-Bruas, N. (2021). Microscopic printing analysis and application for classification of source printer. *Computers & Security, 108*, 102320.

Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.

Vo, P.-Q., Dang, N. T., Nguyen, Q. P., Mai, A., Nguyen, L. T., Nguyen, Q.-T., & Nguyen, N.-T. (2022). Auto machine learning-based approach for source printer identification. In *Recent Challenges in Intelligent Information and Database Systems: 14th Asian*

Conference, ACIIDS 2022, Ho Chi Minh City, Vietnam, November 28-30, 2022, Proceedings (pp. 668–680). Springer.

Xie, S., Girshick, R., Dollár, P., Tu, Z., & He, K. (2017). Aggregated residual transformations for deep neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1492–1500.

Zagoruyko, S., & Komodakis, N. (2016). Wide residual networks. *arXiv preprint arXiv:1605.07146*.

Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. (2017). Mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412*.

Zhang, Z., & Sabuncu, M. (2018). Generalized cross entropy loss for training deep neural networks with noisy labels. *Advances in Neural Information Processing Systems, 31*.