https://doi.org/10.2352/EI.2022.34.10.IPAS-390

DEEP LEARNING BASED UDDER CLASSIFICATION FOR CATTLE TRAITS ANALYSIS

Hina Afridi[1,2] *Mohib Ullah*[1] *Øyvind Nordbø*[2] *Faouzi Alaya Cheikh*[1]

[1] Norwegian University of Science and Technology, 2815 Gjøvik, Norway
[2] GENO SA, 2317 Hamar, Norway.

ABSTRACT

Udder ranking is one of the crucial traits and used extensively in cattle breeding. The analysis of the udder images is challenging due to the variations in the captured conditions of the non-rigid nature of the organ, the farm environment, and disturbances in the form of irrelevant segments of other cattle parts. To this end, we proposed a deep learning-based udder classification algorithm to enhance registrations' precision within cattle breeding. We explore a convolution neural network (CNN), namely the VGG-16 model. The model is trained and validated on a cattle dataset that is collected in Norwegian dairy cattle farms. Expert technicians in the form manually annotate the dataset. We demonstrate that the VGG-16 model used as the backbone can efficiently give an acceptable performance with training and validation accuracy of 97% and 93% respectively on our custom dataset.

***Index Terms*—** Udder classification, Cattle traits, Convolutional neural network, Genetic gain.

1. INTRODUCTION

To improve genetic gain within cattle breeding [1,2], relevant data needs to be collected on a large number of animals. However, for some traits data collection requires a lot of manual work. Thus automation of these processes [3] has the potential to both save costs and improve the genetic gain. Examples of such registration are conformation traits on dairy cows. Currently, this is scored via visual inspection by trained technicians. The international committee for animal recording (ICAR) has developed a list of approved standard traits for conformation recording [4]. These traits should be scored by all organizations in the same way to improve harmonization of conformation traits globally. Furthermore the standard contains a list of five traits (linear, standard, genetic, composite, and general [5]) which are commonly used by organizations in the dairy and dual-purpose breeds worldwide. Linear type traits [6] are the foundation of all systems for describing the dairy cow. Improving linear traits are significantly important for dairy industry for several reasons. They can help in determining individual cow conformation, improving the genetic breeding, and finding characteristics of profitable cows that have longevity and high milk yield.

The total merit index (TMI) [8] describes the relative importance of different trait groupings [7]. Specifically, in the Norwegian Red dairy cattle, the udder conformation traits are contributing with about 20% of the total merit index. Therefore, to have a meaningful basis for predicting udder conformation traits, it is important to filter out the images not containing udder as a pre-processing step. Some of the most important udder traits are udder depth, teat placement, rear udder, udder cleft, fore udder, and udder balance as shown in Fig. 1. For example, good udder depth represents moderate depth relative to the hock with adequate capacity and clearance [9]. A good teat placement trait is squarely placed under each quarter, plumb and properly spaced from side and rear views. Rear udder should be wide and high, firmly attached with uniform width from top to bottom and slightly rounded to udder floor [10]. Good udder cleft is an evidence of a strong suspensory ligament indicated by adequately defined halving. Fore udder of productive cattle is firmly attached with moderate length and ample capacity. Teats should be cylindrical shape and of uniform size with medium length and diameter. Udder balance and texture of good cattle should exhibit an udder floor that is level as viewed from the side [5]. Quarters should be evenly balanced; soft, pliable and well collapsed after milking. It is worth noticing that most of such traits have an optimum. For example, intermediate label is good considering both udder depth and teat length [11]. These traits are recorded manually by the technicians.

In this paper, we explore a deep learning model for udder classification. In a nutshell, the contributions of the paper are the following:

1. We collected cattle udder data at the Norwegian cattle form and manually annotated the data for training a deep model.

2. We demonstrated that a convolution neural network named VGG-16 trained on udder data from scratch gives acceptable performance for the dairy industry.

The current literature does not explore properly the use of deep learning based methods for udder classification. However, some methods have been recently published to analyze udder traits using classical imaging and more recently, deep learning based approaches. For example, Nye et al. [12] introduced a composite deep learning based method to estimate

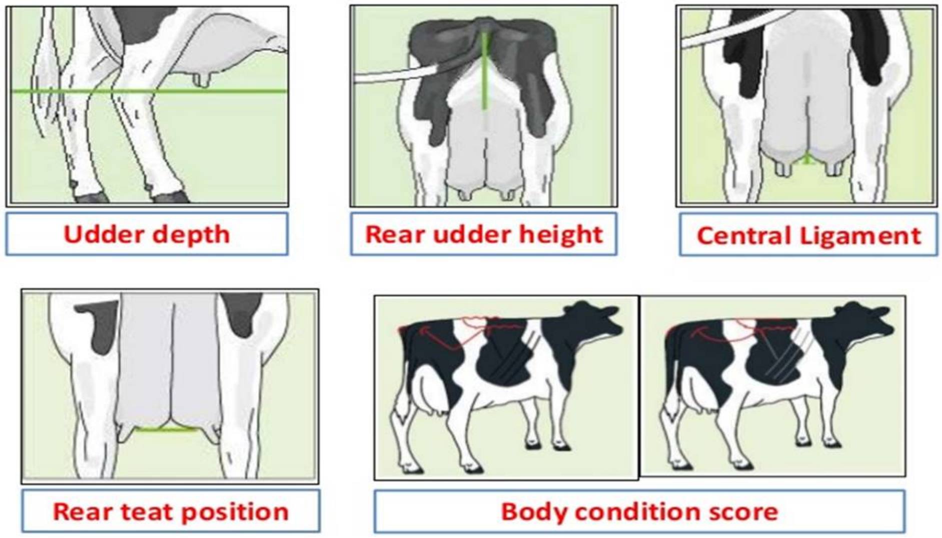

Fig. 1: Different udder traits [7] are depicted including udder depth, rear udder height, central ligament, rear teat position, and body condition score.

conformational traits in dairy cattle. They automatically extracted useful phenotypic information for 14 morphological features. Using pedigree and image information, they estimated high heritability, indicating that meaningful biological information has been extracted automatically from imaging data. Porter et al. [13] used transfer learning to fine-tune Inception Net [14] for mastitis analysis in Holstein cattle. In total, they collected 398 images from 2 commercial farm and tuned the parameters of [14] with 75% of the data. Using classical image processing, Qian et al. [15] come up with a tools to make dairy cow type linear appraisal convenient, swift, and being able to replace manual appraisal. Simiarly, Getu et al. [16] found that the functional conformation traits that influence or facilitate the longevity and reproduction status of dairy cows are the appearance of udder conformation, feet and leg conformation, thoracic and abdominal body conformation, and rump and loin structure. Using 3D scanning technology, Salau et al. [17] presented the concept of a 3D cow scanning by combining the fields of view of six Kinect cameras. Their motivation is to remove if possible, the influence of human operators from conformation recording by gathering data on linear descriptive traits using image analysis. Poppe et al. [18] explored the geometric features of udder and proposed that cartesian teat coordinates measured by automatic milking systems (AMS) provide new opportunities to record udder conformation traits and to study changes in udder conformation genetically and phenotypically within and between parities. They estimated heritability and repeatability of AMS-based udder conformation traits within parities, genetic correlations between parities for AMS-based udder conformation traits, genetic correlations between AMS-based udder conformation traits and classifier-based udder conformation traits, longevity, and udder health. Likewise, Uribe et al. [19] introduced a model to estimate breeding values considering eight conformation traits. These traits include rear udder width, rear udder height, udder depth, and fore udder attachment of cows. They found that additive genetic merit for conformation traits changed with the age of the animal. They found that the single trait, single record model and the simple repeatability model were not appropriate in predicting breeding values at mature ages for rear udder width and rear udder height. More recently, Sinha et al. [20] used principal component analysis to investigate 17 linear udder type traits representing udder and teat conformation and to identify those components having strongest relationship with milk production traits. Correspondingly, Soeharsono et al. [21] predicted daily milk production based on linear body and udder morphometry of Holstein Friesian (HF) dairy cows. We organize the rest of this paper as follows. Section 2 describes our proposed method in details. We provide the datasets description, the experimental results, and discussion in Section 3. Finally, we present the conclusion in Section 4.

2. DEEP BASED UDDER CLASSIFICATION

Deep learning based approaches have presented outstanding performance in many application areas, including computer vision [22], information theory [23], natural language processing [24], and more recently animal breeding [25, 26]. Due to their hierarchical nature, such models have substantial generalization capabilities, especially when trained on proper data considering a specific problem before hand. Inspired by their hierarchical structures, it is viable to explore deep neural networks (DNNs) for the development of automatic udder classification system, considering the particular challenges related to this problem. It is worth noticing that no application based on such approach have yet made it into industrial use.

In this work, we present a deep learning based method for udder classification. In fact, deep neural networks are trainable multi-layer architectures consisting of various feature-extraction phases, followed by a fully connected classification layer [27]. Deep neural networks essentially composed of many layers, and their architectures can be feed-forward or recurrent, having different types of layers and activation functions, and the training can be performed through different optimization techniques. A deep neural network can be modeled from various combinations of fully connected, convolutional, maxpooling, or recurrent layers. Due to their internal architecture and deep nature, they are often trained on large data for extended time, and in general are able to present lower generalization errors.

A deep learning model efficiently extracts and learns multi-level features from the input. The most common type of DNN is the convolution neural network (CNN). CNN is essentially inspired by biological processes in that the connectivity pattern between functions resembles the organization of the animal visual cortex. Furthermore, a convolution neural network [14] [28] [29] is a feed-forward network compose of only convolution layers, maxpooling layers, and fully connected layers. A CNN [30] is the type of deep neural network which is generally used to analyze visual information. In the convolutional layers, a CNN extracts features from the input image in a hierarchical way by exploiting multiple filters. Each filter is made of a set of weights, which are iteratively updated and optimized using an optimization technique. These filters are used to perform a convolution operation on an input image to create a feature map that characterizes the presence of detected features in the input image. The CNN learns the filter coefficients during training depending on the nature of the specific task, and exploits maxpooling layers to sub sample the output. This process propagates the dominant pixels to the next layer in the architecture, dropping the rest.

We explore the VGG-16 model [28] which is one of the most popular deep convolution neural network and commonly used as the backbone architecture for feature extraction in many computer vision problems. It consists of 16 layers where each layers learns different abstract features. The level of abstraction increases from start to the end where earlier layers low level features like edges, colors, arcs and high level layer learns class specific abstract features. In the training stage, we feed the udder and non-udder images to the model. Therefore, during the training stage, the VGG-16 model learns these input images and their corresponding features. In the VGG-16 model, the layers are permitted to be heterogeneous and to deviate widely from biologically inspired connectionist models, for the sake of efficiency, trainability and understandability. There are three types of layers in the VGG-16 model, namely convolutional layers, max-pooling layers, and fully connected layers. A convolutional layer uses a convolution operation to multiply a set of weights with the input. The purpose of the max-pooling layer is to progressively reduce the spatial size of the feature maps to reduce the amount of parameters and computations in the network. A fully connected layer takes the output of the previous layers and flattens them into a single vector as an input for the next step. In total, there are 13 convolution layers, 5 max-pooling layers, and 3 fully connected layers in the VGG-16 model. Only convolutional and fully connected layers contain trainable parameters.

We depict our method in Fig. 2, where all the blue rectangles represent the convolution layers along with the non-linear activation function which is a rectified linear unit (ReLU). The red rectangles represent the max-pooling layers and the circles represent the fully connected layers. The total number of layers having trainable parameters is 16 of which 13 are for convolution layers and 3 for fully connected layers. The last fully connected layer performs the binary classification in the form of udder or non-udder images. To optimize the parameters of the VGG-16 on our dataset, we used the binary cross-entropy cost function. Our binary cross-entropy function is defined as:

$$= -[t \log(p) + (1-t)\log(1-p)] \tag{1}$$

Where t is the true label (either 0 or 1) and p is the probability given by the network. Ideally, the value of p should coincide with t but practically, it is always between 0 and 1. We calculate the binary cross-entropy as the average cross-entropy across all training images defined as:

$$L = -\frac{1}{N}[\sum_{j=1}^{N}[t_j log(p_j) + (1-t_j) log(1-p_j)]] \tag{2}$$

Where t_i is the truth value taking a value 0 or 1 and p_i is the corresponding softmax probability for the i^{th} class. The VGG-16 network transforms the original image via it's layering processing from the original pixel values to the final class scores.

3. EXPERIMENTAL RESULTS

When it comes to the training of the VGG-16 model, it is worth noticing that there is no publicly available udder classification dataset. Therefore, in our work, we train and validate our method using a dataset that we have carefully collected ourselves, with the support of experts in dairy farm. The dataset consists of 232 images for training and 62 images for validation. All images are taken vertically from floor view. Microsoft kinect is used to acquire the images but for the analysis, only the RGB data is used. The resolution of

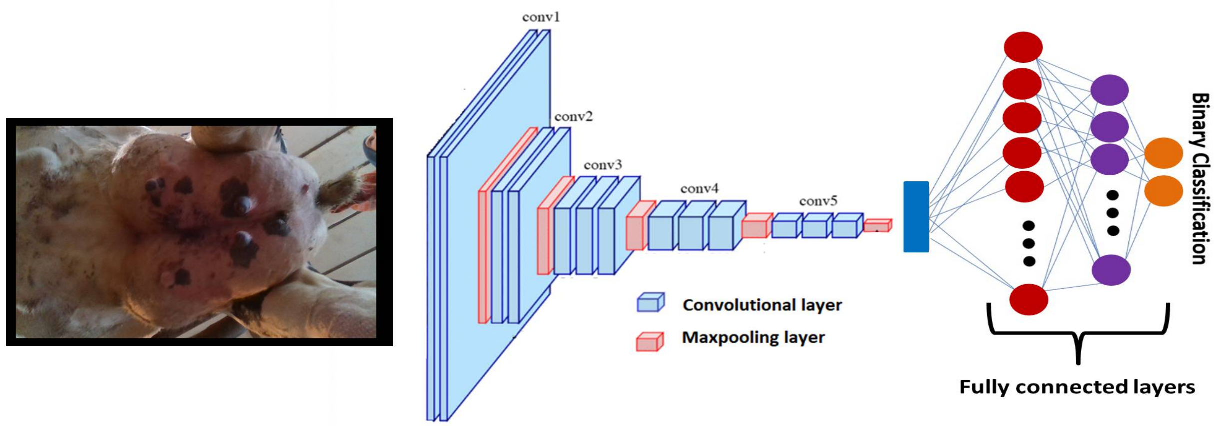

Fig. 2: We feed udder and non-udder images to the network during training. The network consists of convolutional, maxpooling, and fully connected layers. There are 13 convolutional layers and three fully connected layers.

the captures images is 1920x1080. We rescaled the images to 150x150 at the input layer for faster processing. In Fig. 3, sample images for udder and non-udder are shown. The top row shows three sample images representing udders and the bottom row shows three sample images representing non-udder images.

We present the training and validation accuracies in Fig. 4. The red and green graphs show training and validation accuracy, respectively. Both graphs show stable accuracy over time. The best validation accuracy we obtained with the validation data is 93%. We also present the training and validation losses in Fig. 5 by the red and green graphs, respectively. Both validation and training losses converge after epoch number 50. Therefore, the model properly learns without facing the problems of overfitting or underfitting. In Table 1, we report the validation and training accuracies by taking into account different values for the hyperparameters: learning rate and momentum. As can be seen, the variations in the accuracy are not significant by changing the values of these hyperparameters. Which shows the robustness of our method.

Table 1: We report validation and training accuracies by taking into account different values for the hyperparameters: learning rate and momentum.

No	Learning rate	Momentum	Validation accuracy	Training accuracy
1	0.1	0.2	93	97
2	0.2	0.3	92	96
3	0.3	0.4	93	97
4	0.4	0.5	91	95
4	0.6	0.6	92	96
5	0.8	0.7	91	95

4. CONCLUSION

In this paper, we explored a deep learning based model for udder classification using the VGG-16. Our method shows acceptable performance when trained with udder and non-udder images. The reason is that our method learns useful information from both types of images during the training stage. Therefore, our method can be extended to several other applications involving cattle traits analysis. It is important to mention here that our method has been tested considering only udder and non-udder classification. The situation could get challenging when we have more classes in the form of different cattle traits. Therefore, in our future work, we would like to extend our work to multi-class classification for other cattle traits related applications. We will also take into account different augmentation techniques in case we face the problem of the unavailability of sufficient amount of data to train the model.

Acknowledgments

We would like to thank the Research Council of Norway for funding this study, within the BIONÆR program, project number 282252 and the Industrial PhD program, project number 310239. We would also like to thank Hans Snerting and Jan Atle Bakkenget {hans.snerting, jan.atle.bakkenget@tine.no} (both advisors in Tine SA), for the image acquisition.

Fig. 3: Udder and non-udder images. The top row shows three sample images representing udders and the bottom row shows three samples of non-udder images.

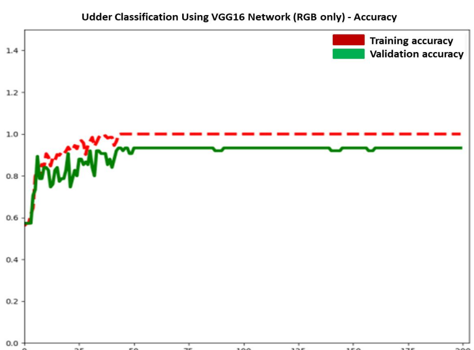

Fig. 4: Training and validation accuracies. The red and green graphs show training and validation accuracies, respectively. Both the graphs show stable accuracies over time.

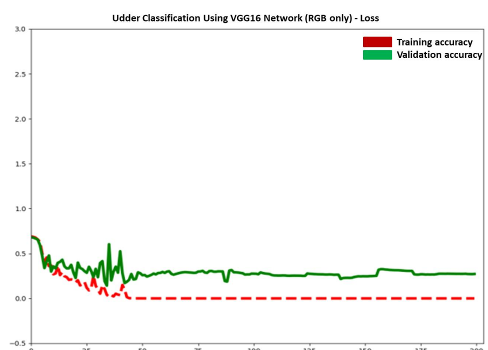

Fig. 5: Training and validation losses. The red and green graphs show training and validation losses, respectively. Both the graphs show stable losses over time.

5. REFERENCES

[1] Anna Wirth, Jürgen Duda, and Ottmar Distl, “Genetic diversity and the impact of the breed proportions of us brown swiss in german brown cattle,” *Animals*, vol. 11, no. 1, pp. 152, 2021.

[2] John B Cole, Sophie AE Eaglen, Christian Maltecca, Han A Mulder, and Jennie E Pryce, “The future of phenomics in dairy cattle breeding,” *Animal Frontiers*, vol. 10, no. 2, pp. 37–44, 2020.

[3] Saleh Shahinfar, Hawlader A Al-Mamun, Byoungho Park, Sidong Kim, and Cedric Gondro, “Prediction of marbling score and carcass traits in korean hanwoo beef cattle using machine learning methods and synthetic minority oversampling technique,” *Meat science*, vol. 161, pp. 107997, 2020.

[4] “https://www.icar.org/index.php/icar-recording-guidelines/,” .

[5] The Global Standard for Livestock Data, “Icar guidelines for conformation recording,” Jun 2018.

[6] Satpal Dahiya, Sunil Kumar, and Manoj Kumar, “Current status of research on linear type traits in indian cattle and future strategies,” *Tropical Animal Health and Production*, vol. 52, pp. 2221–2232, 2020.

[7] “https://www.norwegianred.com/start/norwegian-red/about-norwegian-red/norwegian-red-total-merit-index/,” .

[8] M Kargo, L Hjortø, M Toivonen, JA Eriksson, GP Aamand, and J Pedersen, “Economic basis for the nordic total merit

index," *Journal of dairy science*, vol. 97, no. 12, pp. 7879–7888, 2014.

[9] "https://en.wikipedia.org/wiki/dairy-cattle-evaluation," .

[10] Marc Teissier, Helene Larroque, and Christele Robert-Granie, "Accuracy of genomic evaluation with weighted single-step genomic best linear unbiased prediction for milk production traits, udder type traits, and somatic cell scores in french dairy goats," *Journal of dairy science*, vol. 102, no. 4, pp. 3142–3154, 2019.

[11] Jennifer Salau, Jan H Haas, Wolfgang Junge, and Georg Thaller, "Automated calculation of udder depth and rear leg angle in holstein friensian cows using a multi kinect cow scanning system," *Biosystems engineering*, vol. 160, pp. 154–169, 2017.

[12] Jessica Nye, Laura M Zingaretti, and Miguel Perez-Enciso, "Estimating conformational traits in dairy cattle with deepaps: a two-step deep learning automated phenotyping and segmentation approach," *Frontiers in Genetics*, vol. 11, pp. 513, 2020.

[13] IR Porter, M Wieland, and PS Basran, "Feasibility of the use of deep learning classification of teat-end condition in holstein cattle," *Journal of Dairy Science*, vol. 104, no. 4, pp. 4529–4536, 2021.

[14] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[15] Dongping Qian, Wendi Wang, Xiaojing Huo, and Juan Tang, "Study on linear appraisal of dairy cow's conformation based on image processing," in *International Conference on Computer and Computing Technologies in Agriculture*. Springer, 2007, pp. 303–311.

[16] Addis Getu, Shire Campus, et al., "The role of conformational traits on dairy cattle production and their longevities," *Open Access Library Journal*, vol. 2, no. 03, pp. 1, 2015.

[17] Jennifer Salau, Jan H Haas, Wolfgang Junge, and Georg Thaller, "Automated calculation of udder depth and rear leg angle in holstein-friesian cows using a multi-kinect cow scanning system," *Biosystems engineering*, vol. 160, pp. 154–169, 2017.

[18] M Poppe, HA Mulder, BJ Ducro, and G De Jong, "Genetic analysis of udder conformation traits derived from automatic milking system recording in dairy cows," *Journal of dairy science*, vol. 102, no. 2, pp. 1386–1396, 2019.

[19] H Uribe, LR Schaeffer, J Jamrozik, and TJ Lawlor, "Genetic evaluation of dairy cattle for conformation traits using random regression models," *Journal of Animal Breeding and Genetics*, vol. 117, no. 4, pp. 247–259, 2000.

[20] Rebeka Sinha, Beena Sinha, Ragini Kumari, MR Vineeth, Archana Verma, and Ishwar Dayal Gupta, "Principal component analysis of linear udder type traits and their relationship with milk yield and composition in indigenous sahiwal cattle," *Asian-Australasian journal of animal sciences*, 2020.

[21] Soeharsono Soeharsono, Sri Mulyati, Suzanita Utama, Wurlina Wurlina, Pudji Srianto, Tjuk Imam Restiadi, and Imam Mustofa, "Prediction of daily milk production from the linear body and udder morphometry in holstein friesian dairy cows," *Veterinary world*, vol. 13, no. 3, pp. 471, 2020.

[22] Habib Ullah, Sultan Daud Khan, Mohib Ullah, and Faouzi Alaya Cheikh, "Social modeling meets virtual reality: An immersive implication," in *Pattern Recognition. ICPR International Workshops and Challenges: Virtual Event, January 10–15, 2021, Proceedings, Part IV*. Springer International Publishing, 2021, pp. 131–140.

[23] Abdullah Khalili and Hamid Bouchachia, "An information theory approach to aesthetic assessment of visual patterns," *Entropy*, vol. 23, no. 2, pp. 153, 2021.

[24] Robert Stewart and Sumithra Velupillai, "Applied natural language processing in mental health big data," *Neuropsychopharmacology*, vol. 46, no. 1, pp. 252, 2021.

[25] José Crossa, Johannes WR Martini, Daniel Gianola, Paulino Pérez-Rodríguez, Diego Jarquin, Philomin Juliana, Osval Montesinos-López, and Jaime Cuevas, "Deep kernel and deep learning for genome-based prediction of single traits in multienvironment breeding trials," *Frontiers in genetics*, vol. 10, pp. 1168, 2019.

[26] Akif Quddus Khan, Salman Khan, Mohib Ullah, and Faouzi Alaya Cheikh, "A bottom-up approach for pig skeleton extraction using rgb data," in *International Conference on Image and Signal Processing*. Springer, 2020, pp. 54–61.

[27] Haohan Wang and Bhiksha Raj, "A survey: Time travel in deep learning space: An introduction to deep learning models and how deep learning models evolved from the initial ideas," *arXiv preprint arXiv:1510.04781*, 2015.

[28] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[29] Kyong Hwan Jin, Michael T McCann, Emmanuel Froustey, and Michael Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.

[30] Soonil Kwon et al., "A cnn-assisted enhanced audio signal processing for speech emotion recognition," *Sensors*, vol. 20, no. 1, pp. 183, 2020.