MindPixels: An EEG Image Database

Gerasimos Markantonatos
Department of Electrical and Computer Engineering
National Technical University
Athens, Greece
gerasmark2018@gmail.com

***Abstract*—In the realm of EEG analysis, the complex nature of EEG signals often necessitates their transformation into visual formats, thereby facilitating more accessible data analysis. Despite this practice, there is a notable scarcity of datasets available for such transformed EEG signals. Addressing this gap, we introduce 'MindPixels,' a database developed by transforming EEG signals into visual representations using Gramian Angular Fields. "MindPixels" stands out as a preprocessed, easily accessible repository of EEG-derived image data, effectively streamlining the traditionally time-intensive preprocessing steps. We showcase the utility of "MindPixels" by deploying deep learning techniques for the classification of focal and non-focal EEG GASF images. Our comprehensive evaluation of the "MindPixels" database yields promising outcomes, reinforcing its potential as an invaluable asset in augmenting the interpretation of EEG signals.**

***Index Terms*—Electroencephalogram; Imaging Time Series data; Gramian Angular Field; Spectrogram; Recurrence Plot; Deep Learning;**

I. Introduction

Electroencephalography (EEG), the measure of the electrical fields produced by the active brain, is a neuroimaging technique widely used in research involving neural engineering, neuroscience, and biomedical engineering [1]. Specifically, EEG picks up the electric potential differences, on the order of tens of μV, that reach the scalp when tiny excitatory post-synaptic potentials produced by pyramidal neurons in the cortical layers of the brain sum together. The potentials measured therefore reflect neuronal activity and can be used to study a wide array of brain processes. Thanks to the great speed at which electric fields propagate, EEG has an excellent temporal resolution: events occurring at millisecond timescales can typically be captured. However, EEG suffers from low spatial resolution, as the electric fields generated by the brain are smeared by the tissues, such as the skull, situated between the sources and the sensors. As a result, EEG channels are often highly correlated spatially.

EEG signals are intrinsically noisy and suffer from channel crosstalk. In typical scalp EEG recording settings, each EEG electrode picks up signals from the area nearby, making the spatial resolution coarse (several centimeters). Unmixing the signals is not trivial because of the anisotropic volume conduction characteristics in human brain tissues, skull, scalp, and hair. Therefore, one of the standing challenges in EEG data analysis is how to formulate inputs.

Recently, the availability of large EEG data sets and advances in machine learning have both led to the deployment of deep learning architectures, in the analysis of EEG signals [2]. Three types of input formulation categories have been used for training deep networks: calculated features (41%), images (20%), and the signal values (39%). Notably, many neural networks, especially CNN's, use spectrograms generated from the EEG data as inputs. Other formulations include Gramian Angular Fields and Recurrence Plots.

Spectrograms, being a popular choice, capture the frequency content of EEG signals over time, allowing for the extraction of detailed spectral features. This method is particularly valuable in identifying frequency-specific patterns associated with different cognitive processes or neurological conditions. Gramian Angular Fields, on the other hand, transform EEG time series data into images that highlight temporal dependencies and patterns [3]. This approach contributes to the understanding of dynamic relationships between neural events, enhancing the interpretability of EEG data. Recurrence Plots, another innovative technique, visualize the recurrence of EEG signal patterns, providing insights into the temporal evolution and stability of brain activity [4]. This can be crucial for identifying recurring patterns associated with specific tasks or cognitive states.

Inspired by the success of imaging EEG signals to improve classification in various EEG tasks, we introduce a novel EEG image database called "MindPixels", a collection of EEG signals encoded as images using Gramian Angular Fields.

II. Related Work

A. Unique Contribution: Absence of Existing EEG Image Databases

Before delving into the related work on time series conversion, it is crucial to highlight the groundbreaking nature of "MindPixels" in filling a critical void in the field. Notably, there is a conspicuous absence of dedicated EEG image databases in existing literature. While time series conversion has been a prevalent preprocessing step, the creation of "MindPixels" signifies a departure from this norm, presenting an innovative and direct approach to represent EEG signals as images.

B. EEG Signals to Image Conversion

In prior research, the transformation of EEG signals to image representations has served as a preprocessing step. Spectrograms have been a widely adopted transformation method. Vilamala et al. [5] used spectograms as inputs to a CNN network for sleep stage scoring. Battisti et al. [6] applied spectrograms as a key transformation method in the domain of emotion classification. Bashivan et al. [7] demonstrated the effectiveness of spectrogram-based features in discerning different cognitive states. This commonality in using spectrograms underscores their versatility as a preprocessing technique for various EEG applications.

An intriguing alternative to spectrograms is the use of Gramian Angular Fields (GAF), first introduced by Wang and Oates [3] as a method for imaging time series. Building on Wang's pioneering work, Thanaraj et al. [8] implemented Gramian Angular Summation Fields (GASF) to transform time-series signals into RGB images for epilepsy diagnosis. This innovative approach leverages the rich structure encoded in GASF representations to enhance the interpretability of EEG data, particularly in the context of medical diagnostics. Further demonstrating the versatility of GAF, Ko et al. [9] applied GAF for image conversion, achieving notably higher classification accuracy in schizophrenia diagnosis compared to previous studies. This highlights the potential of GAF as a powerful tool for feature extraction in EEG analysis, particularly in clinical settings. The utility of GAF is not limited to medical applications. Bragin et al. [10] revealed the potential use of GAF conversion for motion imagery recognition in EEG signals. This exploration into diverse applications showcases the adaptability of GAF as a representation method, making it a compelling choice for capturing complex temporal patterns in EEG data across various domains.

Broadening the landscape of image representations for EEG signals, Recurrence Plots (RP) emerge as a distinctive and promising approach. Hatami et al. [11] paved the way by utilizing RPs to transform time-series into 2D texture images, capitalizing on the potential of deep convolutional neural network (CNN) classifiers. This paradigm has since inspired diverse applications, such as Ko et al.'s [9] implementation of Recurrence Plots for converting EEGs into images, particularly for schizophrenia diagnosis. The versatility of RP is further underscored by Paulo et al.'s [12] work, demonstrating its effectiveness in image encoding of EEG signals for drowsiness detection. Additionally, Ravi et al. [13] shed light on the interclass variability captured by recurrence plot images of EEG signals, showcasing their utility in training CNNs for epileptic seizure detection.

III. Methodology

In our methodology, we leverage Gramian Angular Fields to transform electroencephalogram (EEG) signals into a visual format. Furthermore, we propose that Spectrograms and Recurrence Plots may serve as effective alternative methods for representing EEG signals.

A. Gramian Angular Field

Wang and Oates [3] introduced Gramian Angular Summation Field (GASF) as a method to transform time-series signals into images. The encoding process initiates by normalizing the input time series data within the range of [-1, 1]. Subsequently, the normalized or scaled time-series signal undergoes a conversion from Cartesian coordinates to polar coordinates. This transformation retains the temporal information inherent in the input signal. The signal is further warped in the transform domain. In this transformed space, each time point in polar coordinates undergoes a pairwise comparison with every other point to establish temporal correlation. Utilizing the trigonometric cosine function, this comparison results in the Gramian matrix with dimensions [n, n], where n represents the number of sample points in the EEG time epoch. Let $T = \{t_1, t_2, ..., t_n\}$ denote a signal with n samples, and T is rescaled to the interval [-1,1], achieved by the following transformation:

$$\overset{\sim i}{T_0} = \frac{t_i - \min(T)}{\max(T) - \min(T)},$$

where $\min(T)$ and $\max(T)$ represent the minimum and maximum values in the signal, respectively.

The angle ϕ is computed using the equation:

$$\phi = \arccos\left(\overset{\sim i}{T_0}\right).$$

Temporal correlations between adjacent points (i, j) are then computed by summing the angles, resulting in the Gram matrix known as Gramian Angular Summation Field:

$$GASF = [\cos(\phi_i + \phi_j)]$$

This technique enables the transformation of a selected time-series sample into an image. For more detailed information on GASF, refer to [3].

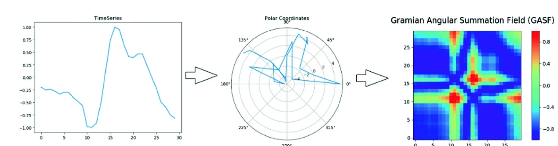


Fig. 1: Illustration of the proposed encoding map of Gramian Angular Fields

B. Spectrogram

A spectrogram is a visual representation of the spectrum of frequencies of a signal as it varies with time. It is built from a sequence of spectra by stacking them together in time and by compressing the amplitude axis into a 'contour map' drawn in a grey scale. The final graph has time along the horizontal axis, frequency along the vertical axis, and the amplitude of the signal at any given time and frequency is shown as a color or intensity.

C. Recurrence Plot

A recurrence plot is a visual tool to inspect the periodic nature of a trajectory through a phase space. It is a technique of nonlinear data analysis through a square matrix, in which the matrix elements correspond to those times at which a state of a dynamical system recurs (columns and rows correspond then to a certain pair of times). It shows for each moment i in time, the times at which a phase space trajectory visits roughly the same area in the phase space as at time j [4]. It can be formulated as:

$$R_{i,j} = \theta(\epsilon - (\|s_i - s_j\|), \quad s(.) \in \mathbb{R}^m, \quad i, j = 1, \ldots, K$$

where K is the number of considered states $s(.)$, ϵ is a threshold distance, $\|.\|$ a norm and $\theta(.)$ the Heaviside function.

IV. Dataset

Our initiative to create an EEG Image Database commences with the utilization of the benchmark Bern-Barcelona (Bern) EEG database [14]. We chose this dataset because it is publicly available for use and provides extensive collection of EEG recordings.

The Bern-Barcelona EEG database is a repository of EEG recordings from both healthy individuals and patients with epilepsy. It encompasses an impressive 3750 single-channel EEG signals, categorized into focal and non-focal types. This rich compilation provides a substantial and diverse pool of data, ideal for our intended transformations and analyses.

For the ”MindPixels” database, we curated a subset of the Bern-Barcelona dataset. Specifically, we selected 50 focal and 50 non-focal EEG recordings, representing a balanced cross-section of the data. Each EEG signal within this subset is methodically segmented, dividing the entire range of 10240 samples (with a sampling rate of 512Hz) into smaller, manageable time segments of 256 samples. This segmentation is a critical step, as it allows for the conversion of these segments into 256x256 pixel-sized images using the Gramian Angular Summation Field (GASF) technique. Consequently, each signal is transformed into 40 distinct GASF images. In total, this process yields 2000 focal GASF images and 2000 non-focal GASF images, derived from 512,000 time samples. This methodical approach is inspired by [8].

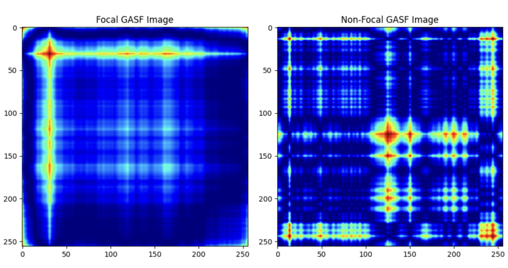


Fig. 2: Representation of Focal and Non-Focal EEG signal using GASF after data splitting

V. Classification

To evaluate the ”MindPixels” dataset’s capability in distinguishing between Focal and Non-Focal signals—a critical task in epilepsy detection—we implemented a suite of Deep Neural Networks. The primary objective of this endeavor was to establish a robust benchmark for the dataset, ensuring that our findings are not only insightful but also easily verifiable and reproducible by the broader scientific community.

A. Image Preparation

To prepare the GASF images for input, we applied a pseudocoloring technique to convert the single-channel images into RGB format. This step is crucial as it allows the pre-trained networks, which are typically designed to process 3-channel color images, to accept our GASF images without significant architectural modifications.

B. Preserving Data Distribution: Stratified Split Approach

In orchestrating our model’s training and validation process, we employed a stratified train-validation split. Our dataset contains 2000 focal and 2000 non-focal images. We obtain 1600 focal and non-focal images for training the deep neural networks and 400 focal and 400 non-focal images for validation. This approach was adopted to preserve the proportional representation of focal and non-focal images across both subsets. By ensuring that each subset maintains the same percentage of focal and non-focal images as the original dataset, we safeguard against potential biases and ensure that the models’ performance accurately reflects their ability to discern between the two classes.

C. Transfer Learning-based CNN

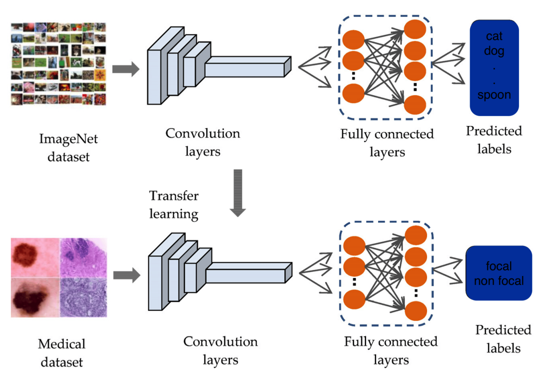
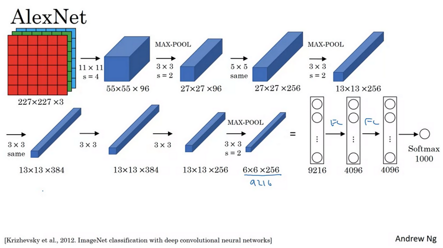


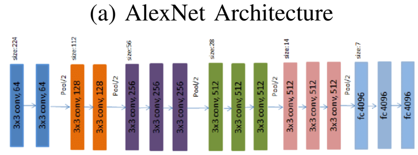
Fig. 3: Transfer Learning

We employed a transfer learning approach with several state-of-the-art CNN architectures, namely AlexNet [15], VGG16 [16], and ResNet [17]. Transfer learning allows us to leverage pre-trained models, which have been trained on large datasets like ImageNet, to extract rich feature representations from our GASF images. This approach is particularly beneficial given the complex nature of EEG data and the need for substantial computational resources to train deep networks from scratch.

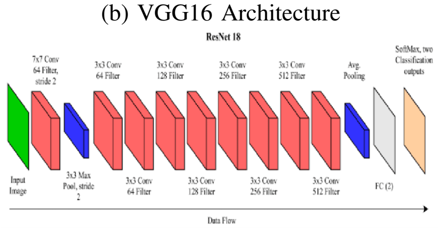
In our approach to optimizing the Convolutional Neural Network (CNN) models for the specific task of distinguishing between focal and non-focal EEG signals, we implemented a targeted fine-tuning strategy. We began by freezing all the pre-trained layers of the CNN architectures, except for the topmost layer. This deliberate preservation of the lower layers ensures that the rich, general feature representations learned from extensive and diverse datasets like ImageNet are maintained. Subsequently, we modified the architecture of the top layer to align with our binary classification task, changing the output to specifically represent the focal and non-focal classes. We then trained our models for 6 epochs. This restrained approach is sufficient due to the pre-trained nature of the networks and helps prevent overfitting, ensuring that the models generalize well to new, unseen data.



(a) AlexNet Architecture



(b) VGG16 Architecture



(c) ResNet Architecture

Fig. 4: Schematic representations of the CNN architectures employed in our study.

D. Custom CNN Model for Epilepsy Detection

In our exploration of CNN architectures, we turn our attention to a model detailed in the work of Thanaraj et al. [8]. Designed for the binary classification of normal and focal EEG signals, this model's architecture and methodology present a compelling foundation for our study. We aim to explore and validate its efficacy within the context of our "MindPixels" dataset. The model consists of three convolutional layers, each employing the ReLU activation function for non-linearity, followed by Batch Normalization (BN) to enhance training stability and performance. Max pooling layers succeed the convolutional blocks to reduce spatial dimensions and highlight predominant features. The network then progresses through Dense layers with Sigmoid activation functions, culminating in a SoftMax output layer for binary classification. For ease of reference and clarity, we provide a figure of the model's architecture as it is proposed in the paper [8].

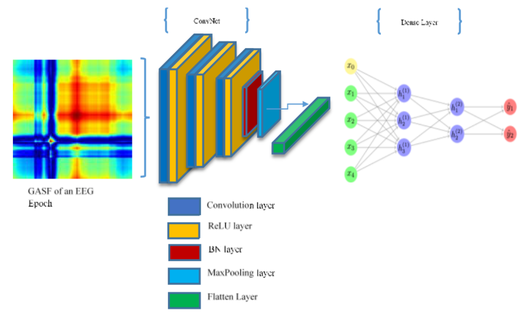


Fig. 5: Custom CNN Architecture

E. Implementation Details

In the development of all aforementioned models, we consistently employed the Cross Entropy Loss function. Optimization was performed utilizing the Adam optimizer. The learning rate was set at $\lambda = 0.0001$, a value determined empirically. Furthermore, we standardized the batch size to 16 across all training procedures to ensure a manageable computational load and used a single NVIDIA TESLA P100 GPU. Our fine-tuned CNN models are trained for 6 epochs and our Custom CNN model is trained for 50 epochs.

VI. Results

To assess the performance of our deep learning architectures, we employ several widely recognized metrics: Precision, Recall, F1 Score, and the Area Under the Receiver Operating Characteristic Curve (AUC). These metrics provide a comprehensive understanding of the models' performance, especially in the context of distinguishing between focal and non-focal EEG signals.

Precision, also known as the positive predictive value, quantifies the number of true positive predictions divided by the total number of positive predictions (both true positives and false positives). It is a measure of a classifier's exactness. A higher precision score indicates a lower false positive rate. The formula for precision is:

$$\text{Precision} = \frac{TP}{TP + FP}$$

where TP represents true positives and FP represents false positives.

Recall, also known as sensitivity, measures the proportion of actual positives that are correctly identified. It quantifies the ability of a model to find all relevant cases within a dataset. The formula for recall is:

$$\text{Recall} = \frac{TP}{TP + FN}$$

where FN represents false negatives. A higher recall score indicates a lower false negative rate.

The F1 Score is the harmonic mean of precision and recall, providing a balance between the two metrics. The F1 Score reaches its best value at 1 (perfect precision and recall) and worst at 0. The formula for the F1 Score is:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The AUC refers to the Area Under the Receiver Operating Characteristic (ROC) Curve. This curve plots the true positive rate (Recall) against the false positive rate (1 - Specificity) at various threshold settings. The AUC is a measure of how well a model can distinguish between classes. An AUC of 0.5 suggests no discrimination ability, while an AUC of 1 indicates perfect discrimination.

The highest classification performance is reported by the Transfer Learning-based VGG16 model with 0.7382 Precision, 0.7825 Recall, F1-Score 0.7597 and 0.7525 AUC. The Transfer Learning-based AlexNet demonstrated similar performance. However, it is worth noting that other models in our study did not achieve the same level of classification performance. The Custom CNN model proposed by Thanaraj et al. [8] exhibited a tendency to overfit the training data. This was evidenced by its high performance during the training phase, yet only mediocre results were observed when the model was applied to the validation data.

TABLE I: Performance Metrics of Deep Learning Architectures

Focal/Non Focal	Precision	Recall	F1-score	AUC
AlexNet	0.7638	0.7275	0.7452	0.7512
VGG16	0.7382	**0.7825**	**0.7597**	**0.7525**
ResNet18	**0.7683**	0.6300	0.6923	0.7200
Custom CNN	0.7179	0.7000	0.7089	0.7125

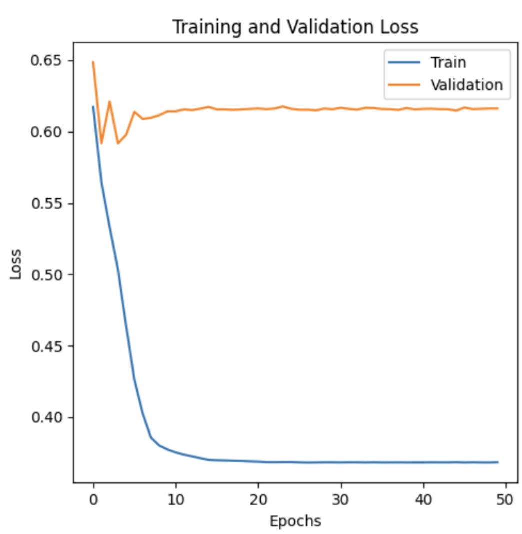


Fig. 6: Train-Validation Loss for Custom CNN

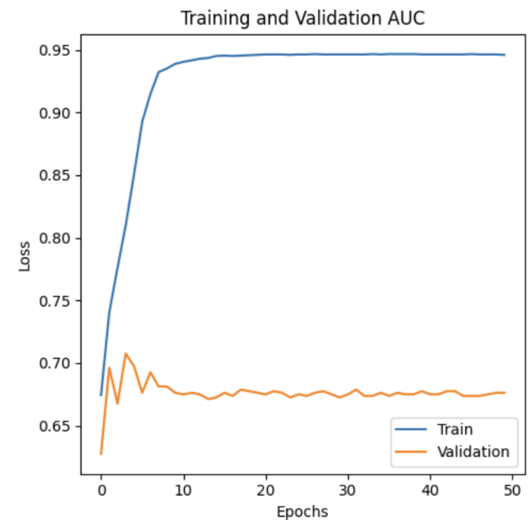


Fig. 7: Train-Validation AUC for Custom CNN

References

[1] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, “Deep learning-based electroencephalography analysis: a systematic review,” *Journal of neural engineering*, vol. 16, no. 5, p. 051001, 2019.

[2] A. Craik, Y. He, and J. L. Contreras-Vidal, “Deep learning for electroencephalogram (EEG) classification tasks: a review,” *Journal of neural engineering*, vol. 16, no. 3, p. 031001, 2019.

[3] Z. Wang and T. Oates, “Imaging time-series to improve classification and imputation,” *arXiv preprint arXiv:1506.00327*, 2015.

[4] J.-P. Eckmann, S. O. Kamphorst, D. Ruelle *et al.*, “Recurrence plots of dynamical systems,” *World Scientific Series on Nonlinear Science Series A*, vol. 16, pp. 441–446, 1995.

[5] A. Vilamala, K. H. Madsen, and L. K. Hansen, “Deep convolutional neural networks for interpretable analysis of EEG sleep stage scoring,” in *2017 IEEE 27th international workshop on machine learning for signal processing (MLSP)*. IEEE, 2017, pp. 1–6.

[6] L. Battisti, A. Ferrato, C. Limongelli, M. Mezzini, and G. Sansonetti, “Deep Learning Based Emotion Classification through EEG Spectrogram Images,” *SOCIALIZE 2023*, 2023.

[7] P. Bashivan, I. Rish, M. Yeasin, and N. C. F. Codella, “Learning Representations from EEG with Deep Recurrent-Convolutional Neural Networks,” *CoRR*, vol. abs/1511.06448, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:17267607

[8] K. P. Thanaraj, B. Parvathavarthini, U. J. Tanik, V. Rajinikanth, S. Kadry, and K. Kamalanand, “Implementation of deep neural networks to classify EEG signals using gramian angular summation field for epilepsy diagnosis,” *arXiv preprint arXiv:2003.04534*, 2020.

[9] D.-W. Ko and J.-J. Yang, “EEG-based Schizophrenia Diagnosis through Time Series Image Conversion and Deep Learning,” *Electronics*, vol. 11, no. 14, p. 2265, 2022.

[10] A. D. Bragin and V. G. Spitsyn, “Electroencephalogram Analysis based on Gramian Angular Field Transformation,” in *CEUR Workshop Proceedings*, vol. 24852019, 2019, pp. 273–275.

[11] N. Hatami, Y. Gavet, and J. Debayle, “Classification of time-series images using deep convolutional neural networks,” in *Tenth international conference on machine vision (ICMV 2017)*, vol. 10696. SPIE, 2018, pp. 242–249.

[12] J. R. Paulo, G. Pires, and U. J. Nunes, “Cross-subject zero calibration driver’s drowsiness detection: Exploring spatiotemporal image encoding of eeg signals for convolutional neural network classification,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 29, pp. 905–915, 2021.

[13] S. Ravi, A. Shahina, N. Ilakiyaselvan, and A. N. Khan, “Epileptic seizure detection using convolutional neural networks and recurrence plots of eeg signals,” *Multimedia Tools and Applications*, vol. 81, no. 5, pp. 6585–6598, 2022.

[14] R. G. Andrzejak, K. Schindler, and C. Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients," *Physical Review E*, vol. 86, no. 4, p. 046206, 2012.
[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.
[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.