



UNIVERSITY OF THESSALY

SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

VOLOS, GREECE

YEAR PUBLISHED: 2024

**Deep Learning Applications in Heart and Brain Health Biometrics
and Brain Tumor Classification**

Diploma Thesis

Marios Andreas Koutras

Representatives from the Examination Committee

Supervisor

Katsaros Dimitrios

Associate Professor, Department of Electrical and Computer Engineering,
University of Thessaly

Member

Thanos Georgios

Laboratory Teaching Staff, Department of Electrical and Computer Engineering,
University of Thessaly

Member

Rafailidis Dimitrios

Associate Professor, Department of Electrical and Computer Engineering,
University of Thessaly

Deep Learning Applications in Heart and Brain Health Biometrics and Brain Tumor Classification

Koutras Marios Andreas

University of Thessaly

Electrical and Computer Engineering

Volos, Greece

kmarios-@uth.gr

Abstract—In recent years, artificial intelligence (AI) has become a prevalent tool in people’s lives. Neural networks have caused a revolution in the medical field, providing powerful tools for disease diagnosis, treatment and its success prediction. In this study we explore the practical use of neural networks in three critical areas of medical care. 1) Brain tumor classification, 2) Prediction of brain age and 3) Arrhythmia diagnosis and arrhythmia type segmentation.

I. INTRODUCTION

1) Automatic segmentation and classification of medical images has a significant role during the diagnostic procedure in areas such as growth prediction, and treatment [1], [2]. An earlier diagnosis leads to a faster treatment which is directly correlated to a higher survival rate. For example in [3] researchers at the Fralin Biomedical Research Institute of Virginia Tech have tackled the problem of chemotherapy resistance of glioblastoma by blocking the signaling protein of PI3K-beta which makes tumor cells more sensitive to temozolomide treatment, a commonly used chemotherapy drug for glioblastoma. [4] found that glioblastoma patients showed that neoadjuvant treatment with an immunotherapy drug almost doubled survival rates, which suggests that the timing of immunotherapy administration can impact patient outcomes, therefore automatic detection/classification procedure is an area worth researching. In the first part of this study we develop a deep convolutional neural network (CNN) used in brain tumor classification. We will be using a combined dataset of multiple smaller datasets and implement the *TumorCNN* model that achieves a **prediction accuracy of 98.78%** and **F1, Recall and Precision scores greater than 97%** in its classification.

2) Chronological age is linked to neurodegenerative conditions such as cognitive impairment and dementia [5]. During the aging process the gray matter of the brain atrophies, along with cortical thinning [6], [7]. Additionally, environmental factors and underlying pathological issues may accelerate brain aging. As a result individual brain aging may vary for each person [8], [9]. Additionally a person’s biological sex has been proven to affect the brain’s volumes [10]–[13] and neurodegenerative changes [14], [15]. These factors can be taken into account by machine learning models that utilize brain magnetic resonance imaging (MRI). BrainAge, also known as the *brain age gap* serves as a potent biometric / biomarker for detecting potential brain related neurodegenerative diseases, brain dysfunction and inter-individual differences in brain aging [16]–[18]. A positive brain age gap indicates accelerated brain aging while a negative brain age gap indicates a slower aging

brain. Structural MRI of the brain can detect neuroanatomical changes that relate to chronological age such as gray matter atrophy. It is therefore established that the chronological age of healthy brain individuals can be accurately estimated using the MRI data [9], [19], [20]. Our proposed *BrainAgeNet* model is an extension of the *TumorCNN* and a fusion of the CNN and MLP models and predicts BrainAge with a **MAE of 11.10 years** and **RMSE of 14.20 years**.

3) Electrocardiograms (ECGs) are widely used by physicians to monitor cardiac health. The challenge posed with the analysis of ECGs is the difficulty of detecting and classifying varying waveforms in a signal. This task is time consuming and error prone for a human. According to data from the Global Burden of Disease Study, NCD Risk Factor Collaboration, NCD Countdown initiative, WHO Global Health Observatory and WHO Global Health Expenditure database, cardiovascular diseases (CVDs) are the cause of death for one third of deaths around the world [21]. Therefore the quick, robust and accurate analysis of cardiac health indicators is highly desirable [22]. In this study we will present the *ECGNet* with a **prediction accuracy of 99.69%** and **F1, Recall and Precision scores greater than 99%** for all classes.

II. LITERATURE REVIEW

Research [23] has proposed the LSFP-Net, a deep unrolled neural network for real-time MRI-guided brain intervention with the goal of enhancing the navigation accuracy and targeting during neurological interventions such as biopsy and deep brain stimulation. The interventional device was a custom-built MR-compatible device with 4 degrees of freedom, 3-DOF ball-joint mechanism for manual trajectory adjustment and 1-DOF lead screw-nut mechanism for remote actuation. Using phantom experiments with a 2D gradient echo sequence with golden-angle radial sampling to evaluate the system’s performance in real-time imaging. The LSFP-Net was employed for real-time MRI reconstruction during the intervention process, exploiting the low-rank background information to enhance the quality of the image. During the phantom experiments / studies, the system displayed high accuracy in the measurement of intervention depth in fruit and porcine-brain phantom experiments with minimal differences between theoretical and measured depths. The LSFP-Net showcases a reconstruction time of 370ms for single group images with a latency of 400ms, making it suitable for real-time MRI applications. The system was also tested for ventricular drainage with promising potential for clinical applications.

[24] presents a deep-learning pipeline for mouse brain mesoscale automatic profiling (D-LMBmap), for the analysis of large imaging datasets generated by whole-brain light-sheet fluorescence microscopy (LSFM) experiments. Compared to other whole-brain connectivity mapping analysis tools, D-LMBmap showcases superior accuracy for certain cell types in all tested modalities and does not require manual annotation for axon segmentation. The researchers presented a qualitative analysis of whole-brain projectome in a single workflow and achieved comprehensive descriptions of neuronal connectivity with high throughput.

In [25], the researchers of the American Medical Informatics Association compared the performance of deep learning versus traditional models for predicting heart failure. According to the study, deep learning models outperformed traditional models in predictive accuracy when analyzing structured and unstructured data from electronic health records for the prediction of heart failure. Traditional models achieved a maximum AUC of 0.75, while deep learning models achieved an AUC of 0.85. The study also found that deep learning models were able to handle missing data more efficiently and provide transparent and interpretable features for better understanding the underlying predictors of heart failure.

[26] researched the feasibility of training CNNs to automate liver segmentation across different imaging modalities and techniques used in clinical practice and apply this technique to enable automation of liver biometry. A 2D U-Net CNN was trained for liver segmentation in two stages by using 330 abdominal MRI and CT examinations. The neural network was trained with unenhanced multiecho spoiled gradient-echo images from 300 MRI examinations to yield multiple signal weightings. Transfer learning was used to generalize the CNN with additional images from 30 contrast material-enhanced MRI and CT examinations. The performance of the CNN was assessed by using a distinct multi-institutional dataset curated from multiple sources (498 subjects). Segmentation accuracy was evaluated by computing Dice scores. These segmentations were used to compute liver volume from CT and T1-weighted MRI examinations and to estimate hepatic proton density fat fraction (PDFF) from multiecho T2-weighted MRI examinations. Dice scores were 0.94 for CT, 0.95 for T1-weighted MRI, and 0.92 for T2-weighted MRI. Liver volume measured with manual and automated segmentation agreed closely for CT and T1-weighted MRI. Hepatic PDFF measured by the two segmentations also agreed closely.

[27] tackles the problem of arrhythmia diagnosis. Arrhythmia diagnosis is a challenge for clinicians, especially in remote and underdeveloped areas. They formulate a time series classification problem, and propose the Shanghai Jiao Tong University Electrocardiogram Network (SJTU-ECGNet). Its key components include CNNs and Residual blocks used for automatic feature extraction from ECG signals. Residual blocks handle short term dependencies while CNN blocks handle spatial features. Bi-LSTM layers, capture long-term dependent features focusing on the rhythm changes between QRS waves. Attention Layers enhance important features and indicates the model's focus on specific input features. Their approach achieves an F1 score of 83.51% and an average AUC ROC score of 0.977 and 93.74% mean accuracy for six common arrhythmias.

III. DATASETS AND FEATURES OF THIS RESEARCH

A. Brain Tumor Classification

The Brain Tumor MRI Dataset [28] employed in this part of the research is a combination of three smaller datasets, the Figshare and Sartaj brain tumor datasets and the Br35H dataset for healthy brain MRIs. It contains a total of 7023 images. The dataset contains glioblastoma, meningioma, pituitary tumor T1-weighted images in the common views (sagittal, coronal and axial). After selecting the appropriate images from the three datasets, the final dataset contains 1621 glioblastoma, 1645 meningioma, 1757 pituitary and 2000 healthy images. Figure 1 displays some instances of the dataset in different sagittal, coronal and axial MRI views. The tumor borders have been highlighted in red.

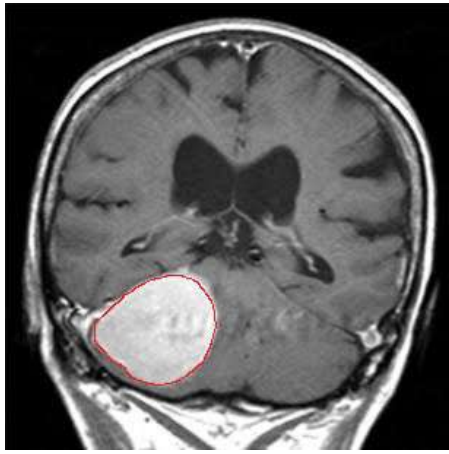
The dataset is minimally processed. Initially all images are converted into a simpler format (JPEG). The results of performing the classification on the original NIFTI / DICOM files does not differ with the classification results of the same files converted to the JPEG format. Therefore, to reduce overfitting and for computational efficiency we will use the JPEG format for this dataset. Each image consists of three input channels (red, green, blue). The images are separated into their respective tumor types (glioblastoma, meningioma, pituitary, healthy), and during the training process they are resized into an image shape of 256x256 to ensure consistency for the model, flipped horizontally at random times to reduce training bias and normalized according to their mean and standard deviation to ensure consistent input range and numerical stability for the model. Finally, 20% of the dataset will be kept for testing, grouped by each category.

B. BrainAge Prediction

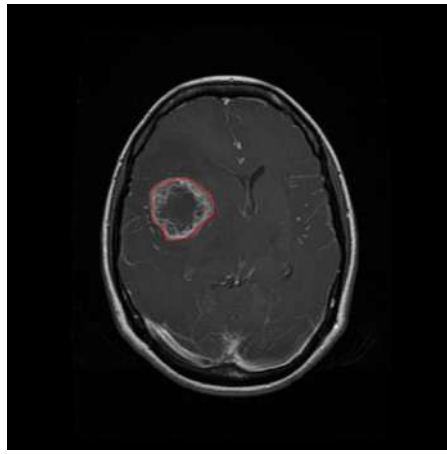
The Information eXtraction from Images (IXI) dataset [29] is a collection of 619 T1 and T2 weighted MRI images from normal, healthy subjects. The data has been collected from three different hospitals in London. The *Hammersmith Hospital* using a Philips 3T system, the *Guy's Hospital* using a Philips 1.5T system and the *Institute of Psychiatry* using a GE 1.5T system.

The MRI files undergo normalization, smoothing, denoising and resampling. The MRI images have differing intensity values depending on the MRI machine and settings used to take the image. Normalization ensures consistency across images. Smoothing is performed by applying a Gaussian filter on the normalized images. This helps reduce the high-frequency noise and enhance the overall quality of the image. The denoising process is performed using a non-local means denoising algorithm. Finally the resampling step ensures the consistent spatial resolution of the images.

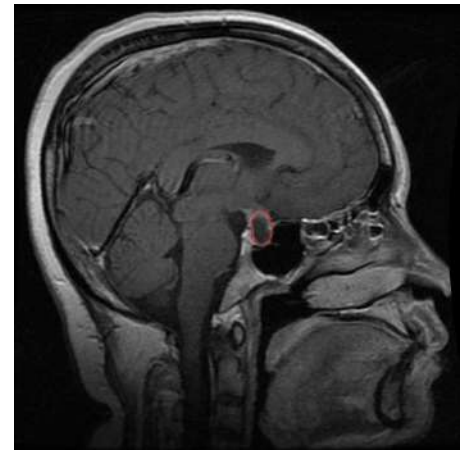
After pre-processing the dataset and removing missing values and broken files the instance count becomes 580 for the T1-weighted images and 575 for the T2-weighted Images. The T1 and T2 weighted images of the dataset are in sagittal and axial view respectively. Each training instance of the dataset is accompanied by the patient's biological sex and the chronological age of the individual in years. The latter serves as the target label of the dataset. Figures 2, 3 display example T1 and T2 weighted images in sagittal and axial view from the



(a) Coronal view of a Meningioma



(b) Axial view of a Glioblastoma



(c) Sagittal view of a Pituitary Tumor

Fig. 1: MRI examples of the Tumor Dataset [28]. The tumor regions have been highlighted in red.

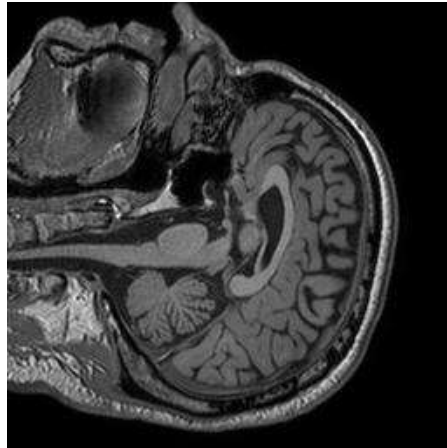
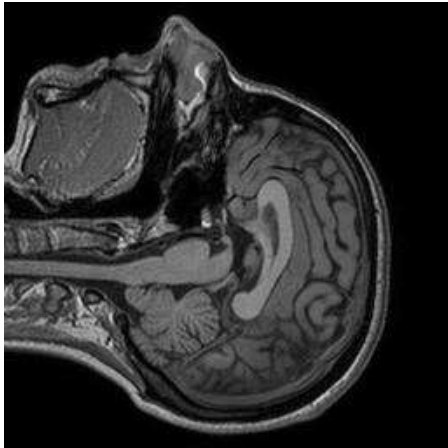


Fig. 2: Sagittal view of random T1-weighted MRIs from the IXI dataset [29].

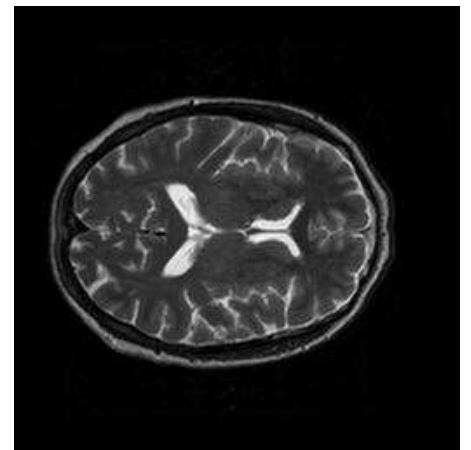
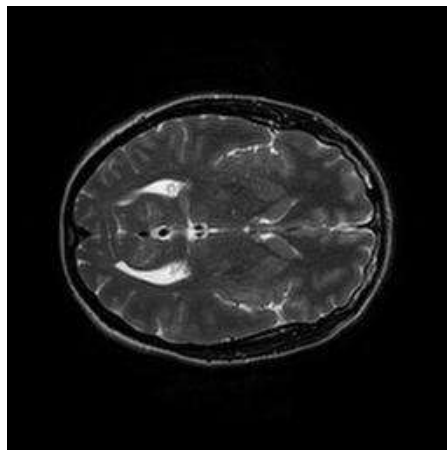
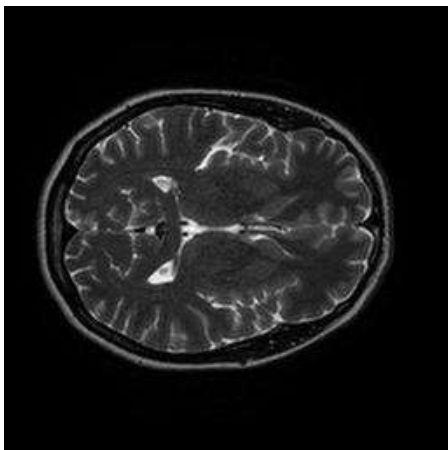


Fig. 3: Axial view of random T2-weighted MRIs from the IXI dataset [29].

IXI dataset. It is noteworthy that T1-weighted MRI enhances fatty tissue, providing crisp images highlighting anatomy and making fluids dark, while T2-weighted MRI enhances the

signal of water which is useful for visualizing inflammation, edema, lesions, etc. Each MRI type can be exploited differently by AI models.

C. Arrhythmia Type Classification

The MIT-BIH dataset [30], is a product of the Boston's Beth Israel Hospital. It contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, from 47 subjects studied by the BIH Arrhythmia Laboratory. Twenty-three recordings were chosen at random from a set of 4000 24-hour ambulatory ECG recordings collected from a mixed population of inpatients (about 60%) and outpatients (about 40%) from the hospital. The 25 remaining recordings were selected from the same set to include less common but clinically significant arrhythmias which would not be well-represented in a small random sample. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range. Two or more cardiologists independently annotated each record; disagreements were resolved to obtain the computer-readable reference annotations for each beat (approximately 110,000 annotations in all) included with the database.

ECG signals collected as raw data in a clinical environment are often affected by noise such as power frequency interference, baseline drift and EMG interference. We can denoise the data to improve the performance of our classification model. Low pass filters, bandpass filters and wavelet transforms [31] are widely used in ECG data processing. We employ the wavelet transform method to pre-process the ECG signals, which is an algorithm that decomposes non-stationary signals into scale signals of different bands. We select the Sym4 function family as the wavelet function [32] and a threshold of 0.04 to cut off any noise.

The classes of the dataset are *N*, *L*, *R*, *A*, *V*. The ECG signals are segmented and pre-processed, with each segment representing a heartbeat. The '*N*' class represents the normal heartbeat (cardiac complex) indicating no abnormalities in the heart rhythm. The '*L*' and '*R*' indicate the presence of a left (right) bundle branch block. '*A*' represents the premature atrial beat or premature atrial contraction (PAC) and finally the ventricular premature beat or premature ventricular contraction (PVC) is labeled as '*V*'. We will briefly explain some examples of data from the ECG dataset used in this research. Fig. 4 illustrates the most important segments of an ECG.

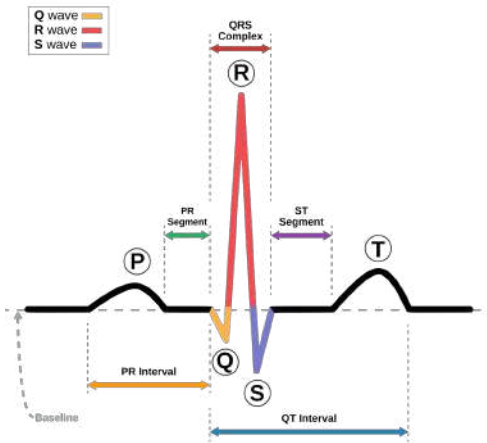


Fig. 4: Illustration of a heartbeat viewed from an ECG.

- P-wave: The Atrial Contraction, meaning the contraction of the cardiac muscle in the atrium which increases pressure, causing bloodflow.
- PR Interval: Time taken from excitation to spread from the SA node across the atrium and down to the ventricular muscle via the His-Purkinje system.
- QRS Complex: Ventricular Contraction.
- ST Segment: Ventricular Relaxation.
- T-wave: Ventricular Repolarization.

Fig. 5 displays a healthy ECG beat. The P-Q interval and QRS complex are shorter than 120 ms. All waves are normal.

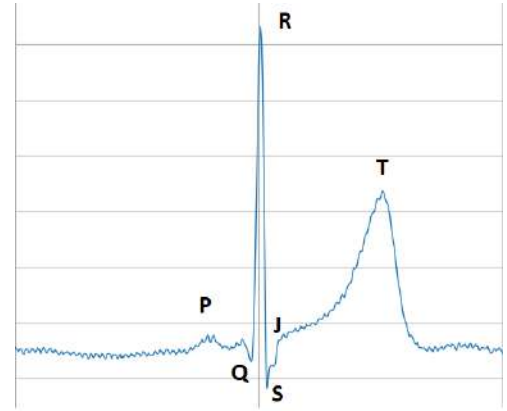


Fig. 5: Illustration of a normal ECG beat from the MIT-BIH [30] dataset.

Fig. 6 displays the V1 ECG of a Right Bundle Branch Block (RBBB). The QRS complex is widened (> 120 ms) (1) and the RsR' pattern in the QRS complex (2) is a typical indicator of RBBB.

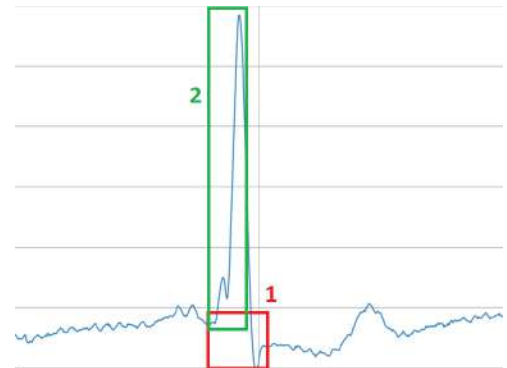


Fig. 6: Illustration of an RBBB from the MIT-BIH [30] dataset.

Fig. 7 displays a Premature Ventricular Contraction (PVC). A P-wave from the SA-Node causes a normal QRS complex (1). This is followed by a wide unnatural ventricular QRS complex (2). We notice ST depression (3) and T-wave inversion (4).

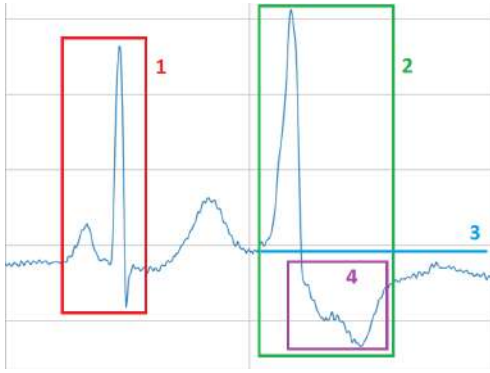


Fig. 7: Illustration of a PVC beat from the MIT-BIH [30] dataset.

IV. MEDICAL TERMINOLOGY

In this section we will present a brief overview on the medical terminologies used in this study and their importance.

A. Glioblastoma

Glioblastoma, also known as glioblastoma multiforme (GBM) is the most commonly appearing malignant brain tumor and represents approximately 57% of all gliomas and 48% of all primary malignant central nervous system (CNS) tumors [33]. Patients with this disease usually showcase a poor prognosis with a median survival of less than 2 years. Glioblastoma typically appears in the cerebral hemispheres, in the frontal, temporal and parietal lobe. The parietal lobe is involved in the processing of sensory information which impairs the neurological function of the body and reduces the quality of life for patients, caregivers and families alike [34].

B. Meningioma

Meningioma is the most common primary central nervous system tumor accounting for 37.6% of them; and approximately 50% of all benign brain tumors [35]–[37]. Meningiomas appear on the membranes that cover the brain and spinal cord (meningeal layers), known as the meninges. Brain meningiomas can be categorized as convexity, falx, parasagittal, olfactory groove, sphenoid, petrous fossa, posterior fossa, intraventricular and intra-orbital meningiomas. Spinal meningiomas are rare and occur in the thoracic spine, causing back pain [38], [39].

C. Pituitary Tumors

The pituitary gland is located withing the sella turcica, a bony depression in the sphenoid bone. It is divided into the anterior and posterior segments. The anterior pituitary (adenohypophysis) is comprised of glandular tissue while the posterior pituitary (neurohypophysis) is comprised of neural tissue. The pituitary gland is connected to the hypothalamus with the pituitary stalk which has a portal circulation for the anterior pituitary and neuronal connection for the posterior pituitary. The anterior pituitary gland is regulated by hormones secreted by the hypothalamus, while the posterior pituitary hormones are produced by the hypothalamus itself and are subsequently stored in the posterior pituitary for release [40],

[41]. Pituitary tumors in the anterior pituitary gland account for 15% of intracranial neoplasms and are usually benign [42], [43]. In rare scenarios, pituitary tumors could metastasize which are then termed as pituitary carcinoma [44]. Systemic metastasis to the pituitary gland which could originate from the lung, breast, prostate, renal, colon, etc. differs from pituitary carcinoma [45]. The most known complication of pituitary tumors is vision loss [46]. Other complications include diabetes insipidus [47], pituitary apoplexy (hemorrhage in the pituitary gland) [40], meningitis, CSF leak, internal carotid artery rupture [48]–[50].

D. Left (Right) Bundle Branch Block

Left (Right) bundle branch block (LBBB / RBBB) occur when the physiologic electrical conduction system of the heart in the His-Purkinje system, is altered, compromised or interrupted, causing blockage along the pathway that electrical impulses travel to make the heart beat resulting in widened QRS and ECG vector changes [51], [52]. Despite being an indicator of underlying heart disease, as isolated incidents, LBBB and RBBB do not cause clinical concern and can be seen in patients with no clinical disease. However with the appropriate clinical context, LBBB and RBBB can be of significant consequence and importance, especially in patients with acute chest pain, syncope and in those suffering from heart failure with reduced ejection fraction (HFrEF) [53].

LBBB and RBBB ECG criteria are defined by the American College of Cardiology (ACC) and American Heart Association (AHA) as follows.

LBBB [54]

- QRS duration greater than 120 ms
- Rhythm must be of super-ventricular origin
- Lead V1 should have either a QS or a small r wave with large S-wave
- Lead V6 should have a notched R wave and no Q wave

RBBB [55]

- QRS Duration greater than 120 ms
- In leads V1 and V2, there is an RSR' pattern
- In leads I and V6, the S-wave is of greater duration than the R wave, or the S-wave is greater than 40 milliseconds
- In leads V5 and V6, there is a normal R wave peak time
- In lead V1, the R wave peak time is greater than 50 milliseconds

E. Premature Atrial Contraction

Premature atrial contractions (PACs) are contractions of the atria triggered by the atrial myocardium which did not originate from the Sinoatrial Node (SA node). PACs can be caused by a variety of medical diseases, structural abnormalities, pharmaceuticals and non regulated compounds and

are prevalent amongst young and old patients and can show independently of risk factors and previously known medical conditions [56]. Because of the usually benign nature of PACs their pathophysiology has not been well established. Common theories are based on abnormal automaticity of the cardiac myocytes, increased eliciting events from chemical or physical instigators and reentry of retrograde impulse [57]. In an ECG PACs typically have a normal QRS complex and a normal, short or longer PR interval than sinus rhythm. Non conducted PACs can occur which have no QRS complex following the PAC. PACs can arise from one location (unifocal) or different locations (multifocal). The P wave of PACs usually occurs earlier than the sinus P wave with a different morphology and axis from the sinus P wave. Similar to RBBB and LBBB, PACs should be analyzed in conjunction with underlying pathologies. Complications of PACs include, atrial fibrillation [58], ischemic stroke [59], increased cardiac related and all-cause mortality [60].

F. Premature Ventricular Contraction

Premature ventricular contractions (PVCs) occur when the heartbeat is initiated by the purkinje fibers and not the Sinoatrial Node (SA node). Since a PVC occurs before a regular heartbeat there is a pause in the ECG rhythm before the next regular heartbeat [61]. They can present themselves as isolated incidents or in repeated patterns. Two consecutive PVCs are called *doublets*, and three consecutive PVCs are named *triplets*. However, three or more consecutive PVCs are classified as ventricular tachycardia. PVCs are caused by excessive caffeine consumption, excess catecholamines, anxiety, electrolyte abnormalities [62]. PVCs can also directly result in cardiac and non-cardiac pathologies such as cardiomyopathy, mitral valve prolapse and myocardial infection. Additionally, structural heart diseases that result in conduction pathway degeneration due to tissue alterations can cause PVCs [63]. The diagnostic process involves an ECG to look for ectopic ventricular beats. PVCs are infrequent in most patients and therefore a short ECG period will most likely fail to capture PVCs. This also allows to differentiate PVCs with PACs. The ECG should include electrolyte abnormalities which translates to peaked T waves and a prolonged QT, left ventricular hypertrophy, Q waves, loss of R waves and possible RBBB/LBBB and finally, acute ischemia may present with ST-segment elevation/depression and/or T wave inversion. Complications include an increased risk of heart failure and ventricular tachyarrhythmia [64], [65] and cardiomyopathy [66].

V. TERMINOLOGY OF ALGORITHMS AND NEURAL NETWORKS

Convolutional neural networks are feed forward neural networks (FFNN) that can extract features from data with convolution structures. Their architecture is inspired by visual perception [67]. Unlike traditional feature extraction methods [68]–[70], CNNs do not extract features manually. CNN kernels represent different receptors which respond to different features. Activation functions simulate the function that only neural electric signals that exceed a certain threshold will be transmitted to the next neuron. Using loss functions and optimizers the network can learn the expected features. Compared to simpler neural networks, CNN have specific advantages; 1)

Locality. CNNs have the ability to detect local spatial features within the data. 2) Weight sharing. This reduces the number of parameters. 3) Downsampling. The amount of data is reduced while retaining useful information.

A. 2D Convolutional Networks

2D CNNs have been extensively used in real world applications such as Age and gender recognition from speech, bearing fault diagnosis (BFD) for various real world applications, human action recognition from sensor data, hyper-spectral image classification and medical imaging radiology applications [71]–[76]. Fig. 8 displays the basic building blocks of a 2D CNN configuration. This model classifies a single channel image of shape 24×24 into two categories. The network consists of two convolution layers with 4 neurons and two pooling layers with 6 neurons. The output of the final pooling layer is flattened and passed to a fully connected layer and the output layer which produces the classification prediction. Each kernel has a size of (K_x, K_y) . The convolution reduces the image dimensions by $(K_x - 1, K_y - 1)$. The subsampling factors (S_x, S_y) are used in the pooling layers. In Fig. 8 $K_x = K_y = 4$ and $S_x = S_y = 3$ for the first pooling layer and $S_x = S_y = 4$ for the second one.

B. 1D Convolutional Networks

Fig. 9 displays the basic building blocks of an 1D CNN configuration. The conventional 2D CNNs are used exclusively on two dimensional data such as images (MRI images in this study) and videos. The modified version of 2D CNNs, the 1D CNNs have been developed for numerous applications in the fields of ECG, mechanical or aerospace structures, high power circuitry, power engines and motors [77]–[86]. According to these studies, 1D convolutions prove advantageous for certain applications for the following reasons; 1) Lower computational complexity. An image with $N \times N$ dimensions convolved with a $K \times K$ kernel will have a computational complexity of $O(N^2 K^2)$. The corresponding 1D convolution will have a complexity of $O(NK)$. 2) Lower number of parameters. Most 1D CNN applications use compact models consisting of less than 10k parameters, while 2D CNN applications use “deeper” models with millions of parameters. 3) Hardware Limitations. 1D CNNs can be trained on CPU devices due to their low numbers of parameters and computational complexity. 4) Suitable for low-cost real-time applications on mobile and hand-held devices [77]–[89].

C. Training Process

CNNs are trained in a supervised manner using the back-propagation algorithm. During each iteration the gradient magnitude (sensitivity) of each trainable parameter is computed and used to update the CNN parameters iteratively until a certain stopping criterion is reached. The backpropagation algorithm uses certain gradient descent based optimization methods such as Stochastic Gradient Descent (SGD), Momentum SGD [90], AdaGrad [91], RMSProp [92], Adam [93] and its variants [94].

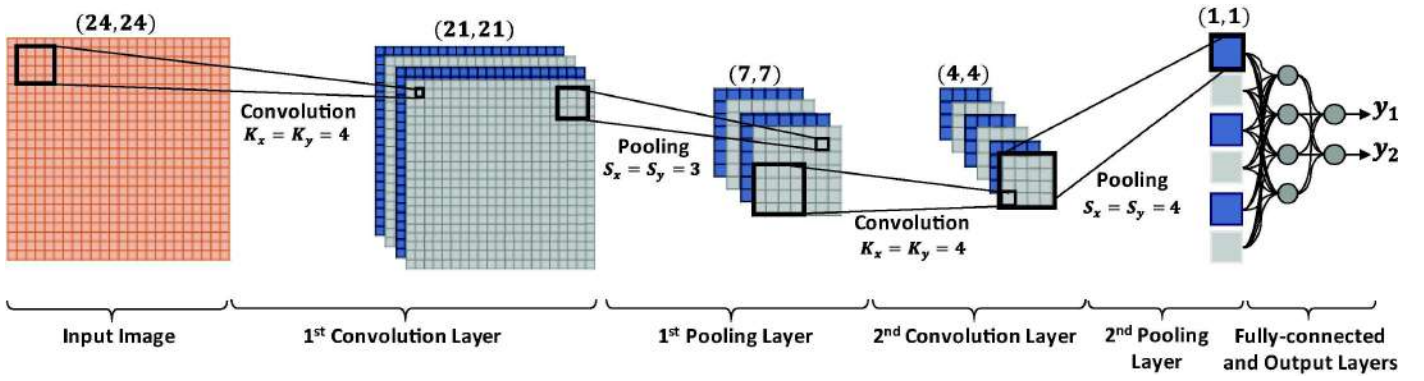


Fig. 8: Illustration of the 2D CNN architecture.

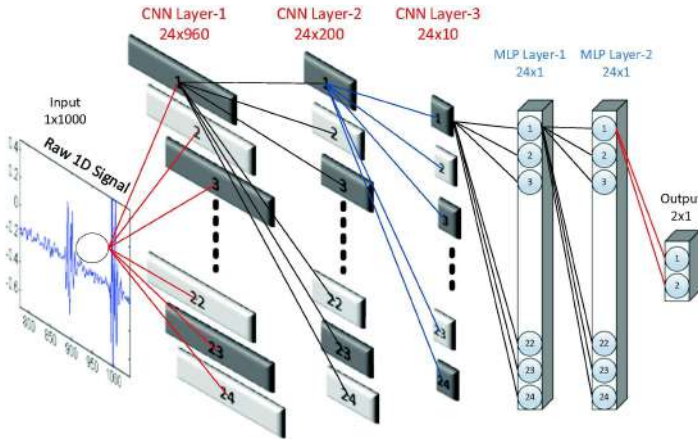


Fig. 9: Illustration of the 1D CNN architecture.

1) *1D CNN Forward and Backpropagation:* In the forward pass of a 1D CNN, we perform convolutions between the input sequence and learnable filters (kernels). Let's denote:

- X as the input sequence with dimensions (N, C, L) , where:
 - N is the batch size,
 - C is the number of input channels,
 - L is the length of the input sequence.
- K as the set of filters with dimensions (F, C, M) , where:
 - F is the number of filters,
 - M is the length of each filter.

The output of the convolution operation can be calculated as:

$$Z_{i,f}^{[l]} = \sum_{c=0}^{C-1} \sum_{m=0}^{M-1} X_{i+m,c}^{[l]} \cdot K_{m,c,f}^{[l]} + b_f^{[l]}$$

where:

- $Z_{i,f}^{[l]}$ is the activation at position i of the f^{th} filter in the l^{th} layer.
- $X_{i,c}^{[l]}$ is the value at position i of the c^{th} channel in the input sequence.

- $K_{m,c,f}^{[l]}$ is the parameter (weight) corresponding to the m^{th} position of the c^{th} channel of the input and the f^{th} filter.
- $b_f^{[l]}$ is the bias term associated with the f^{th} filter in the l^{th} layer.

After obtaining the convolutional outputs, we apply an activation function (e.g., ReLU) element-wise:

$$A_{i,f}^{[l]} = g(Z_{i,f}^{[l]})$$

where $g(\cdot)$ is the activation function.

Optionally, we may include a pooling layer to reduce the sequence length:

$$\text{Pooling operation: } P_{i,f}^{[l]} = \text{PoolingFunction}(A_{i,f}^{[l]})$$

The gradient with respect to the pooling layer is passed back:

$$\frac{\partial \mathcal{L}}{\partial P_{i,f}^{[l]}} \Rightarrow \text{Pooling backward}$$

The gradients with respect to the pre-activation are multiplied element-wise with the gradients of the activation function:

$$\frac{\partial \mathcal{L}}{\partial Z_{i,f}^{[l]}} = \frac{\partial \mathcal{L}}{\partial A_{i,f}^{[l]}} \cdot g'(Z_{i,f}^{[l]})$$

where $g'(\cdot)$ is the derivative of the activation function.

The gradients with respect to the filters $K^{[l]}$ and the input $X^{[l]}$ are computed:

$$\frac{\partial \mathcal{L}}{\partial K_{m,c,f}^{[l]}} = \sum_i X_{i+m,c}^{[l]} \cdot \frac{\partial \mathcal{L}}{\partial Z_{i,f}^{[l]}}$$

$$\frac{\partial \mathcal{L}}{\partial X_{i,c}^{[l]}} = \sum_f \sum_m K_{m,c,f}^{[l]} \cdot \frac{\partial \mathcal{L}}{\partial Z_{i-m,f}^{[l]}}$$

Update the filter weights and biases using the gradients obtained:

$$K_{m,c,f}^{[l]} := K_{m,c,f}^{[l]} - \alpha \frac{\partial \mathcal{L}}{\partial K_{m,c,f}^{[l]}}$$

$$b_f^{[l]} := b_f^{[l]} - \alpha \sum_i \frac{\partial \mathcal{L}}{\partial Z_{i,f}^{[l]}}$$

where α is the learning rate.

2) *2D CNN Forward and Backpropagation:* In the forward pass of a 2D CNN, we perform convolutions between the input image and learnable filters (kernels). Let's denote:

- X as the input image with dimensions (N, C, H, W) , where:
 - N is the batch size,
 - C is the number of input channels,
 - H is the height of the input,
 - W is the width of the input.
- K as the set of filters with dimensions (F, C, HH, WW) , where:
 - F is the number of filters,
 - HH is the height of each filter,
 - WW is the width of each filter.

The output of the convolution operation can be calculated as:

$$Z_{i,j,f}^{[l]} = \sum_{c=0}^{C-1} \sum_{p=0}^{HH-1} \sum_{q=0}^{WW-1} X_{i+p,j+q,c}^{[l]} \cdot K_{p,q,c,f}^{[l]} + b_f^{[l]}$$

where:

- $Z_{i,j,f}^{[l]}$ is the activation at position (i, j) of the f^{th} filter in the l^{th} layer.
- $X_{i,j,c}^{[l]}$ is the pixel at position (i, j) of the c^{th} channel in the input image.
- $K_{p,q,c,f}^{[l]}$ is the parameter (weight) corresponding to the (p, q) position of the c^{th} channel of the input and the f^{th} filter.
- $b_f^{[l]}$ is the bias term associated with the f^{th} filter in the l^{th} layer.

After obtaining the convolutional outputs, we apply an activation function (e.g., ReLU) element-wise:

$$A_{i,j,f}^{[l]} = g(Z_{i,j,f}^{[l]})$$

where $g(\cdot)$ is the activation function.

Optionally, we may include a pooling layer to reduce the spatial dimensions:

$$\text{Pooling operation: } P_{i,j,f}^{[l]} = \text{PoolingFunction}(A_{i,j,f}^{[l]})$$

The gradient with respect to the pooling layer is passed back:

$$\frac{\partial \mathcal{L}}{\partial P_{i,j,f}^{[l]}} \Rightarrow \text{Pooling backward}$$

The gradients with respect to the pre-activation are multiplied element-wise with the gradients of the activation function:

$$\frac{\partial \mathcal{L}}{\partial Z_{i,j,f}^{[l]}} = \frac{\partial \mathcal{L}}{\partial A_{i,j,f}^{[l]}} \cdot g'(Z_{i,j,f}^{[l]})$$

where $g'(\cdot)$ is the derivative of the activation function.

The gradients with respect to the filters $K^{[l]}$ and the input $X^{[l]}$ are computed:

$$\frac{\partial \mathcal{L}}{\partial K_{p,q,c,f}^{[l]}} = \sum_i \sum_j X_{i+p,j+q,c}^{[l]} \cdot \frac{\partial \mathcal{L}}{\partial Z_{i,j,f}^{[l]}}$$

$$\frac{\partial \mathcal{L}}{\partial X_{i,j,c}^{[l]}} = \sum_f \sum_p \sum_q K_{p,q,c,f}^{[l]} \cdot \frac{\partial \mathcal{L}}{\partial Z_{i-p,j-q,f}^{[l]}}$$

Update the filter weights and biases using the gradients obtained:

$$K_{p,q,c,f}^{[l]} := K_{p,q,c,f}^{[l]} - \alpha \frac{\partial \mathcal{L}}{\partial K_{p,q,c,f}^{[l]}}$$

$$b_f^{[l]} := b_f^{[l]} - \alpha \sum_i \sum_j \frac{\partial \mathcal{L}}{\partial Z_{i,j,f}^{[l]}}$$

where α is the learning rate.

D. Loss Functions

1) *Categorical Cross-Entropy:* The Categorical Cross-Entropy function combines both the softmax activation function and the cross-entropy loss into a single function. This method is commonly used for multi-class classification problems.

Softmax

Given an input vector $\mathbf{z} = [z_1, z_2, \dots, z_C]$ (where C is the number of classes), the softmax function outputs a probability distribution over these classes:

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^C e^{z_k}}, \quad \text{for } j = 1, 2, \dots, C$$

where:

- $\sigma(\mathbf{z})_j$ is the probability of the j -th class.
- z_j is the raw score (logit) for the j -th class.

Cross-Entropy Loss

The cross-entropy loss measures the difference between the predicted probability distribution (obtained from the softmax function) and the true distribution. For a single example, if the true label is represented as a one-hot encoded vector $\mathbf{y}$ and the predicted probabilities are $\mathbf{p} = \sigma(\mathbf{z})$, the cross-entropy loss $\mathcal{L}$ is defined as:

$$\mathcal{L}(\mathbf{y}, \mathbf{p}) = - \sum_{j=1}^C y_j \log(p_j)$$

where:

- $\mathbf{y} = [y_1, y_2, \dots, y_C]$ is the one-hot encoded true label vector.
- $\mathbf{p} = [p_1, p_2, \dots, p_C]$ is the predicted probability vector from the softmax function.

Combining Softmax and Cross-Entropy

In practice we combine the softmax function with the cross-entropy loss for numerical stability and computational efficiency. Given the logits $\mathbf{z}$ and the true class label y (where y is a scalar representing the correct class index), the combined loss for a single example is:

$$\mathcal{L}(\mathbf{z}, y) = -\log \left(\frac{e^{z_y}}{\sum_{k=1}^C e^{z_k}} \right)$$

This simplifies to:

$$\mathcal{L}(\mathbf{z}, y) = -z_y + \log \left(\sum_{k=1}^C e^{z_k} \right)$$

where:

- z_y is the logit corresponding to the true class y .
- $\sum_{k=1}^C e^{z_k}$ is the sum of exponentiated logits over all classes.

Batch Loss

For a batch of N examples, the average cross-entropy loss is computed as:

$$\mathcal{L}_{\text{batch}} = \frac{1}{N} \sum_{i=1}^N \mathcal{L}(\mathbf{z}^{(i)}, y^{(i)})$$

where:

- $\mathbf{z}^{(i)}$ is the logits vector for the i -th example.
- $y^{(i)}$ is the true class label for the i -th example.

2) *Mean Square Error Loss*: The Mean Squared Error function computes the mean squared error (MSE) loss between the predicted and the actual values and is widely used for regression tasks, where the goal is to predict continuous values.

E. Mean Squared Error (MSE) Loss

The mean squared error loss measures the average of the squares of the differences between the predicted values and the true values. Given a batch of predictions and true values, the MSE loss is defined as follows:

$$\text{MSE}(\hat{\mathbf{y}}, \mathbf{y}) = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2$$

where:

- N is the number of examples in the batch.
- $\hat{\mathbf{y}} = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_N]$ is the vector of predicted values.
- $\mathbf{y} = [y_1, y_2, \dots, y_N]$ is the vector of true values.
- $\hat{y}_i$ is the predicted value for the i -th example.
- y_i is the true value for the i -th example.

Batch Loss

For a batch of N examples, the MSE loss can be computed as:

$$\mathcal{L}_{\text{batch}} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2$$

This formula calculates the average squared difference between the predicted and true values for the entire batch.

F. Classification Performance Metrics

Accuracy

Accuracy is the ratio of the number of correct predictions to the total number of predictions made. It is defined as:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

Mathematically, if TP , TN , FP , and FN are the number of true positives, true negatives, false positives, and false negatives respectively, accuracy can be expressed as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision

Precision is the ratio of correctly predicted positive observations to the total predicted positives and a measure of the model's exactness. Precision is defined as:

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

Recall

Recall (also known as Sensitivity or True Positive Rate) is the ratio of correctly predicted positive observations to all the observations in the actual class and a measure of the model's completeness. Recall is defined as:

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

F1-Score

The F1-score is the weighted average of Precision and Recall and takes both false positives and false negatives into account. The F1-score is defined as:

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

Substituting the definitions of Precision and Recall, the F1-score can be written as:

$$\text{F1-Score} = \frac{2TP}{2TP + FP + FN}$$

G. Gradient-weighted Class Activation Mapping

The Gradient-weighted Class Activation Mapping (Grad-CAM), an explainable AI technique, extends class attention mapping by taking the gradient of the target class score, with respect to the feature maps of interest (convolution layers) [95], [96]. This provides a finer localization of the important regions by weighting the importance of each feature map, based on its gradient during backpropagation by employing a heatmap [97]. Despite Grad-CAM primarily for classification tasks [96], it has been previously adapted to the problem of regression tasks and specifically into gray matter age prediction [98].

VI. NEURAL NETWORK MODELS

Fig. 10 displays a convolution block which is the basic building block of the neural networks in this research. A convolution block consists of the following components: The Convolution, Batch Normalization, ReLU and the Max Pooling Layer.

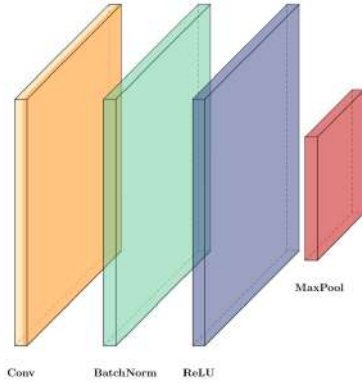


Fig. 10: Breakdown of the Convolution Block

A. TumorCNN - Brain Tumor Classification

The proposed *TumorCNN* is shown in Fig 11. The input of our model are JPEG format images converted directly from NIFTI files. These images are first pre-processed and converted to the same size (256, 256), converted to a tensor, and corrected according to their mean and standard deviation. This is a three channel input passed to the first convolutional block which receives 3 channels and outputs 12, using a kernel size of 3×3 . After undergoing Batch Normalization, ReLU and Max Pooling, this output is passed to the second convolution block (Block2). Block 2 consists of two convolutional layers back to back, one with the input of 12 channels and the output of 32 and the other with an input of 32 and an output of 64 with a kernel of 4×4 . Then the batch normalization, ReLU and pooling take place. After the data passes through the third convolution block, consisting of 256 channels and a 5×5 kernel, it is flattened, starting from the first dimension (not 0) and then enters the MLP layers. Finally, we use the logarithmic softmax activation function in the output layer.

B. BrainAgeNet - BrainAge Prediction

The BrainAgeNet is depicted in Fig 12. The 3D NIFTI file is loaded and pre-processed. We normalize the image data to

values between 0 and 1, since the MRI images in this dataset were collected from different Hospitals with different MRI machines with different parameters. Additionally we resample images with a third dimension different than 150 to have a consistent third dimension of 150 slices for the T1-weighted images and 130 slices for the T2-weighted images, depending on what the input of the model is. The images are reshaped into a format consistent with PyTorch's convolutional networks, which require the channels of the image to be the first dimension. Finally the processed image, the sex of the patient and the age corresponding to that image are converted to tensors. Therefore each instance of the dataset is a minimally processed MRI NIFTI image along with its sex (1 = male, 2 = female) and the age label. The processed image is fed to the first convolutional block (Block1) of Fig. 12 and then propagated through the network. The categorical sex information is fed to the layer *FC_Age1* and propagated through the MLP network. The outputs of the two networks are concatenated along the first axis and then passed through another MLP network which produces the output. Since this is a regression problem, no activation function is used.

C. ECGNet - ECG Classification

The ECGNet in Fig. 13 takes raw signal data as its input reshaped in the form of an array of shape (N, 1). The 1D Convolution is suitable for this type of data since ECGs have significant short term dependencies that can be exploited to predict the short term future behavior of the heart. In this network the convolution blocks are simplified and only the convolution, ReLU and max pooling blocks are used as materials. The batch normalization blocks are not mandatory since the model still showcases excellent performance without them. Note that the presence of the BatchNorm block also impacts the Grad-CAM feature visualization heatmaps. The first convolution layer outputs 16 filters with the input of 1. The second convolutional layer outputs 32, the third 64 and the fourth and final, 128. All max pooling operations are performed with a pooling kernel of size 3 and a stride of 2. The output of the convolutional blocks is flattened and passed to the dense layers, one with 128 neurons, the second with 64 and the third one (output layer) with 5, which is equal to the number of classes. Finally we use the softmax activation function to obtain the probability distribution of the classes in the output layer. This network showcases excellent performance for all categories of arrhythmia.

VII. RESULTS

Below we will present the results of the study and the performance metrics of the model. Each metric is the result of training the model 10 different times and calculating the mean of its performance metrics.

A. TumorCNN - Tumor Classification

We have tested numerous optimization methods mentioned in section V-C. We train the model using the Adam optimizer which achieves significantly better performance and faster convergence than all other methods and the cross-entropy loss which is widely used for classification problems. We use a learning rate and weight decay of 0.0001. These settings are able to converge to the absolute minimum of the loss function

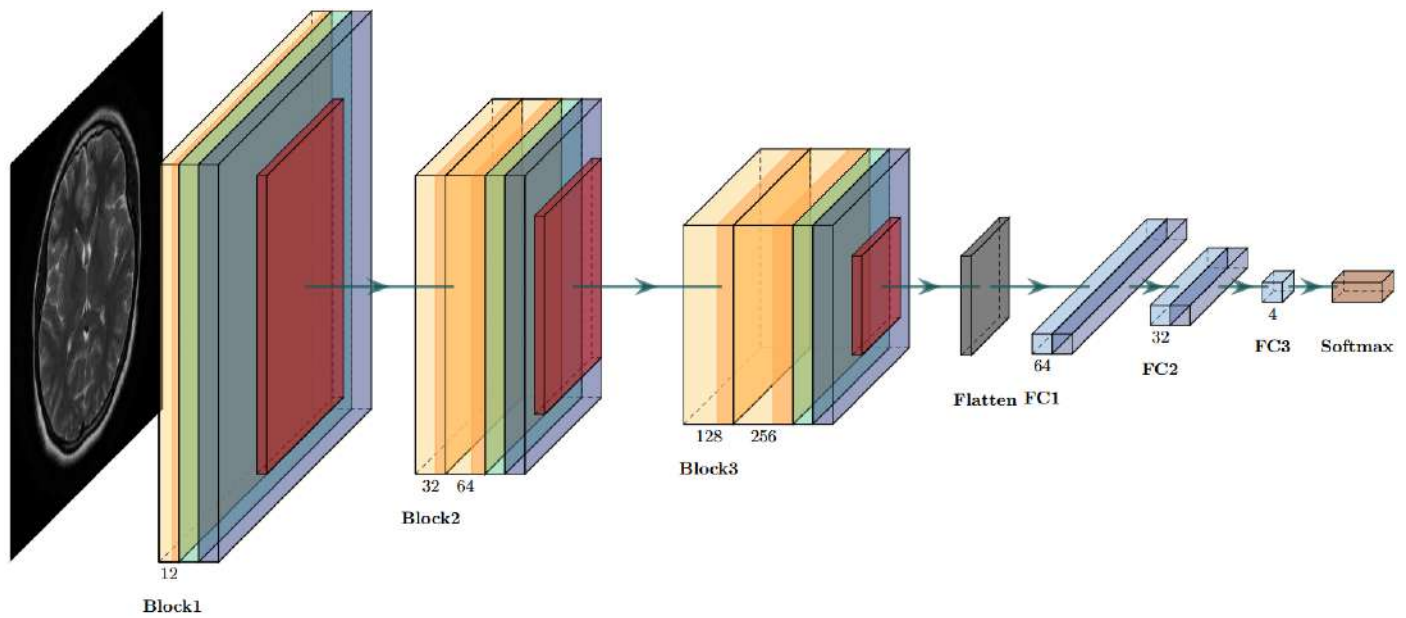


Fig. 11: The **TumorCNN** model architecture.

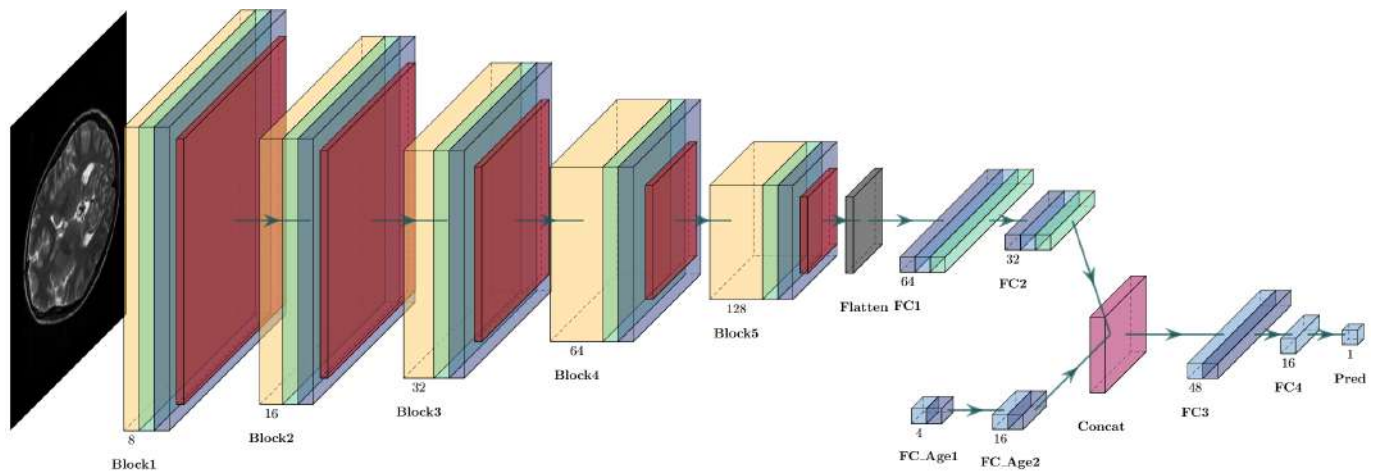


Fig. 12: The **BrainAgeNet** model architecture.

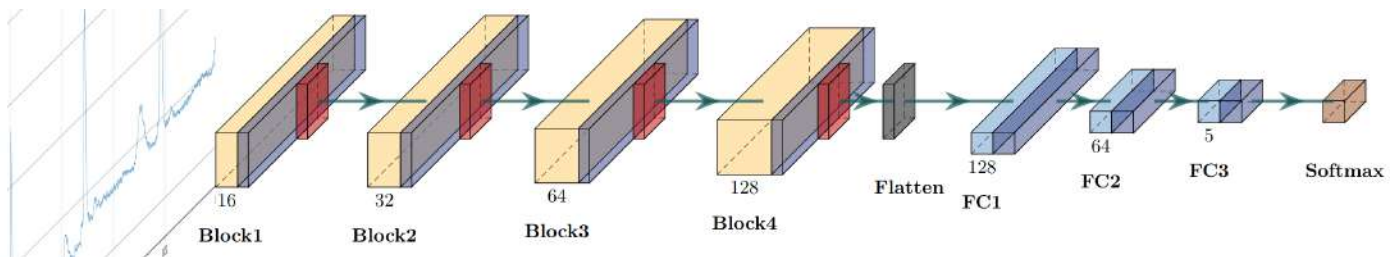


Fig. 13: The **ECGNet** model architecture.

achieving the best results in terms of performance metrics. We train the model for 30-50 epochs which is the number of iterations required for the loss function to reach the minimum with the small learning rate we selected. During each iteration we average the training accuracy and loss over all iterations

and calculate the training accuracy and loss. Fig. 14 displays the cross-entropy losses for 30 training epochs for the training and testing set. The training loss is smoother and close to the mean for all epochs. The testing loss has a greater standard deviation for all epochs, especially during the first few epochs

before the model manages to generalize. The average accuracy per epoch is presented in Fig. 15. Training with a very small learning rate is required to get the best performance on the testing set. Finally the model achieves the performance metrics for each class as shown in Table I.

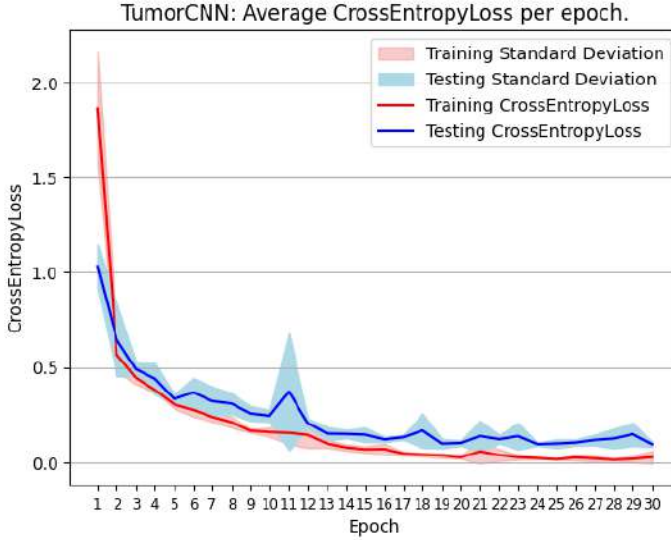


Fig. 14: Training/Testing Loss per epoch for the TumorCNN model.

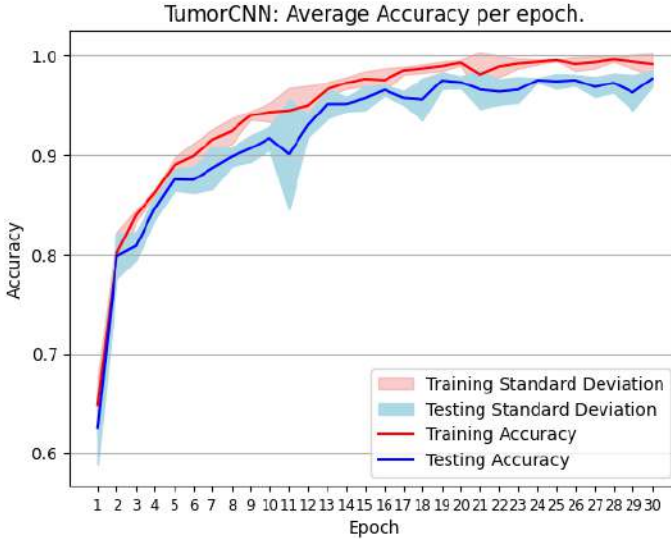


Fig. 15: Training/Testing Accuracy per epoch for the TumorCNN model.

The TumorCNN model demonstrates excellent prediction performance for all classes with an overall accuracy of 98.25% and f1, recall, precision scores $\geq 97\%$. The *Healthy* and *Pituitary* classes achieve almost unit precision, recall and F1. This is attributed to the clear differentiation between healthy brain images which exhibit no structural abnormalities and pituitary tumor images which can be identified by the distinct location (pituitary gland) of the tumor. However Glioblastoma and Meningioma Tumors often require more than one MRI

TABLE I: TumorCNN Classification Performance.

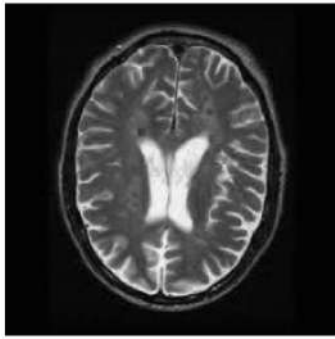
Tumor Type	Precision	Recall	F1-Score
Glioblastoma	1.00	0.98	0.99
Meningioma	0.98	0.97	0.97
Healthy	0.98	1.00	0.99
Pituitary	1.00	1.00	1.00
Accuracy 0.99 (1311 samples)			
Macro avg	0.99	0.98	0.99
Weighted avg	0.99	0.99	0.99

view (axial, coronal, sagittal, see Fig. 1). For example, the axial view of the glioblastoma in Fig. 1 is often not enough to diagnose the glioblastoma tumor type since its specific characteristics may not be fully discernible. If this tumor was a meningioma, it would appear at the meninges, the upper layers of the brain. This is not distinguishable from an axial brain MRI. Therefore a coronal or sagittal view of the brain would be a necessity for an accurate diagnosis. This could be a factor for the slightly lower performance of the model, in predicting glioblastomas and meningiomas (see Table I). In a further attempt to assess the performance of our model we will utilize Grad-CAM to analyze some of the important features of each class and the areas of the MRI that the model "sees". Regions of the image with warmer colors (yellow, orange, red) are more significant to our model's predictions.

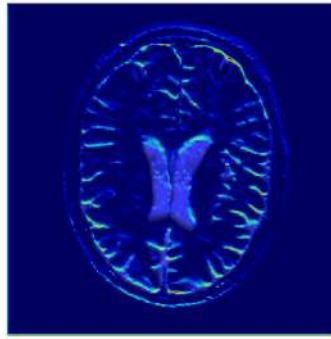
1) *Healthy Brain Feature Visualization*: Fig. 16 displays the important features when a healthy brain MRI is propagated through the network. The first, second and third convolutional layers (Fig. 16b, 16c, 16d) primarily emphasize on the gray matter, near the frontal and temporal lobes, the frontal horn of the lateral ventricle and the choroid plexus. The attention of the fourth and fifth convolutional layers extends on the genu of corpus callosum, anterior cerebral arteries, and the overall area of the temporal lobe (Fig. 16e, 16f). The biggest part of the brain is emphasized which is reasonable in order to deduct that the brain is healthy.

2) *Glioblastoma Feature Visualization*: Fig. 17 shows the model activation for a glioblastoma MRI. In the glioblastoma MRI the first layer (Fig. 17b) exhibits minimal activation within the brain area and heightened activation towards the surrounding black regions encompassing the cranium. This is a phenomenon observed often with the glioblastoma MRIs of the dataset and displays possible bias. That is due to the predominant use of axial views in glioblastoma MRIs, which consequently biases the dataset and the learning process. The second and third layers (Fig. 17c, 17d) display gradient activations in the meninges of the temporal lobe. The final two layers, 4 and 5 (Fig. 17e, 17f) focus on the brain area outside of the tumor region. This means that they exclude the area of the tumor and make their prediction depending of the absence of other types of tumors.

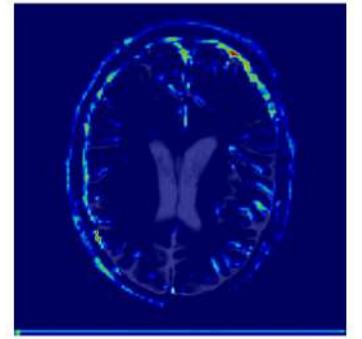
3) *Meningioma Feature Visualization*: Fig. 18 illustrates the convolution outputs for meningioma MRIs. The model effectively distinguishes the tumor regions clearly in every convolutional layer. The first convolution layer (Fig. 18b)



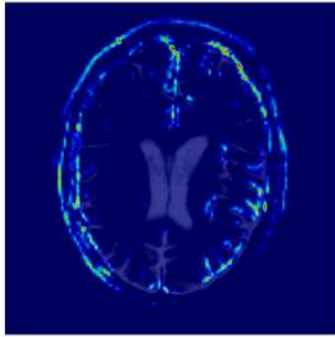
(a) Axial Healthy Brain



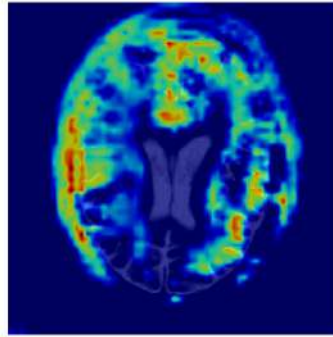
(b) Conv1



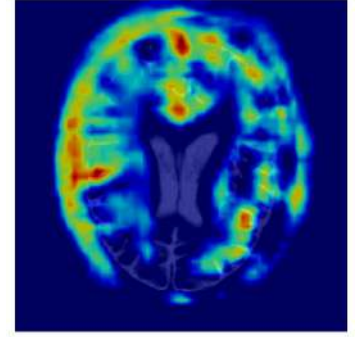
(c) Conv2



(d) Conv3

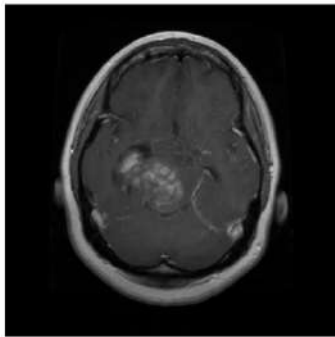


(e) Conv4

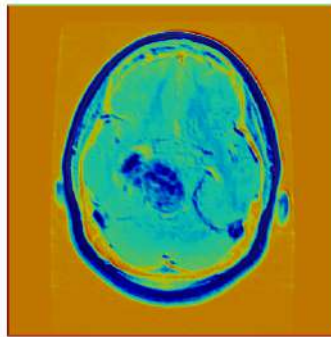


(f) Conv5

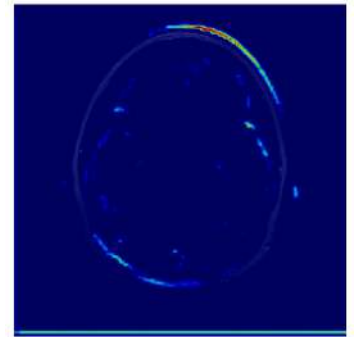
Fig. 16: Healthy brain from the Tumor dataset [28] and its convolutional features.



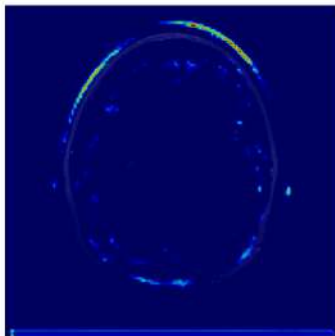
(a) Axial Glioma



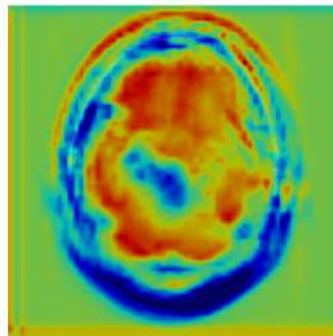
(b) Conv1



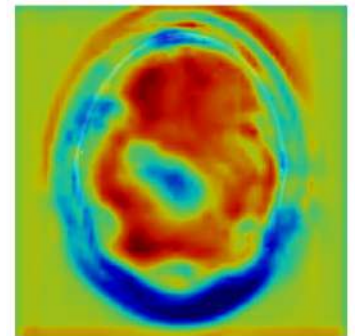
(c) Conv2



(d) Conv3

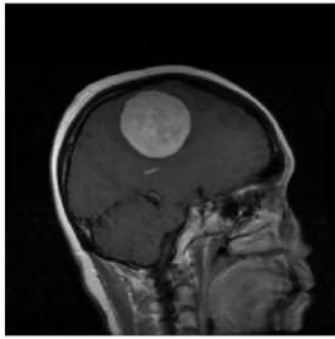


(e) Conv4

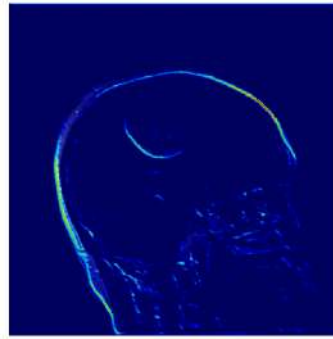


(f) Conv5

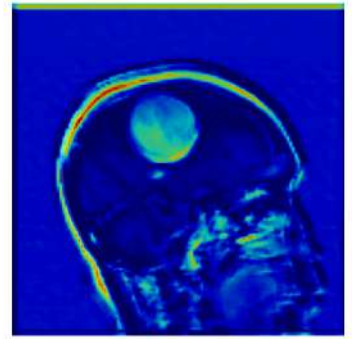
Fig. 17: Axial Glioblastoma from the Tumor dataset [28] and its convolutional features.



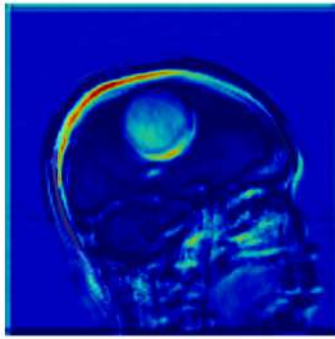
(a) Sagittal Meningioma



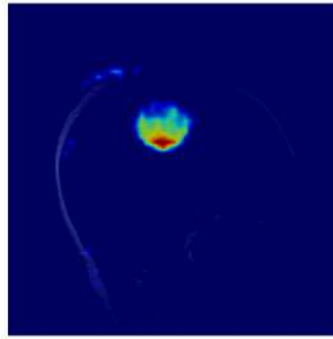
(b) Conv1



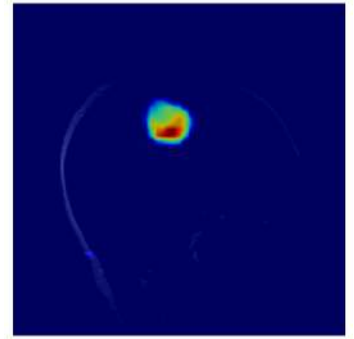
(c) Conv2



(d) Conv3

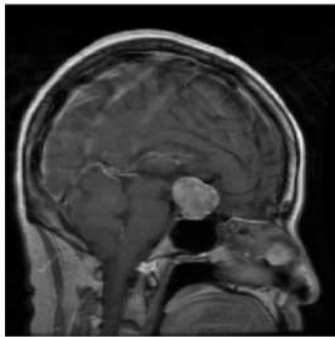


(e) Conv4

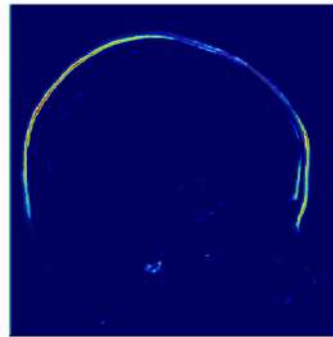


(f) Conv5

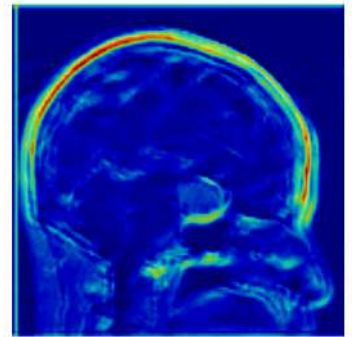
Fig. 18: Sagittal Meningioma from the Tumor dataset [28] and its convolutional features.



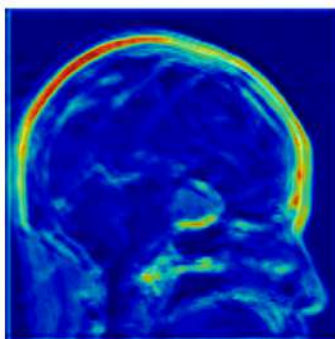
(a) Sagittal Pituitary Tumor



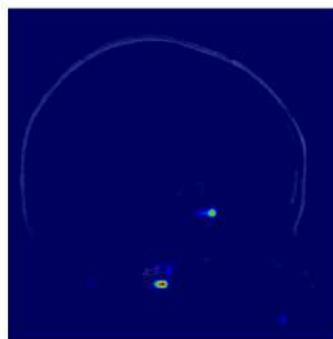
(b) Conv1



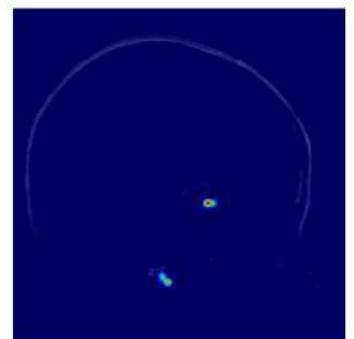
(c) Conv2



(d) Conv3



(e) Conv4



(f) Conv5

Fig. 19: Sagittal Pituitary Tumor from the Tumor dataset [28] and its convolutional features.

appears deactivated since it mostly contributes to the prediction of healthy and glioblastoma brain MRIs. The second and third layers (Fig. 18c, 18d) direct their attention to the meninges and to the borders of the meningioma tumor region. This attention is intensified in the final two layers (4 and 5), which solely focus on the meningioma borders (Fig. 18e, 18f).

4) *Pituitary Tumor Feature Visualization*: Fig. 19 displays the model outputs for a sagittal view of a tumor in the pituitary gland. The first convolutional layer (Fig. 19b) has minimal activation. This is reasonable since it mostly activates for healthy and glioblastoma images and usually focuses on the cranium and meninges. The second (Fig. 19c) and third (Fig. 19d) layers start focusing on the pituitary gland, in the tumor region borders and the structural abnormality of the pituitary that has a growth. The final layers (Fig. 19e, 19f) emphasize their attention on the region of the tumor.

The model demonstrates excellent performance and precise attention regions in most images. However a notable drawback is the bias susceptibility if the dataset is not diverse enough to include all possible MRI views for all possible classes. For instance, if the training dataset predominantly consists of axial images for glioblastomas and sagittal or coronal images for meningiomas, the model may incorrectly associate axial images with glioblastomas and sagittal or coronal images with meningiomas. As a result, the model fails to generalize and correctly focus on the tumor region, which is crucial for its precision and robustness against outliers. This phenomenon can be seen in Fig. 17, where a lot of attention is given to the region outside the cranium.

B. BrainAgeNet - Brain Age Prediction

The model is trained using the Adam optimizer with a learning rate of 0.001 and a weight decay of 0.0001, for 10 epochs and a batch size of 24, since NIFTI files demand a lot of memory. We use the MSELoss function, which is widely used for regression problems for continuous variables such as the human age. BrainAgeNet's average training and testing loss is depicted in Fig. 20. BrainAgeNet displays low variance in its loss for the training set and higher variance for the testing set. 10 epochs are sufficient for model training.

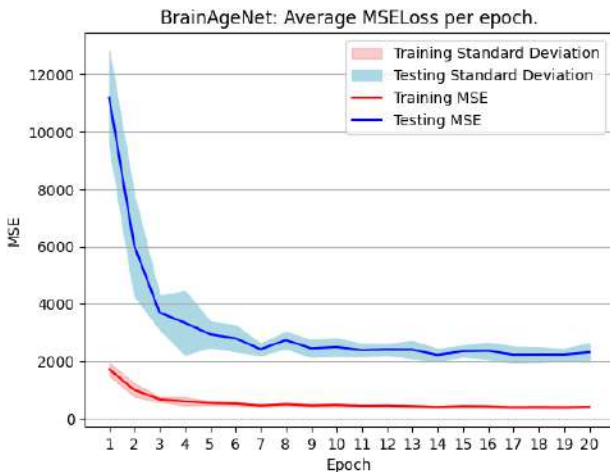


Fig. 20: BrainAgeNet losses for 20 epochs.

The model's convolutional layers gradually display their attention to deeper areas of the brain with each layer. The first layer weakly focuses on the entire image 21b. The second layer focuses on the meninges and the outer shape and structure of the brain while showing attention to the frontal and superior sagittal sinus as well as the shape of the frontal horn of lateral ventricle 21c. The third layer enhances its attention to the same points 21d. The fourth and fifth layers begin focusing on the gray matter and corpus callosum. 21e, 21f. This is expected since gray matter volume and connectivity begins decreasing with age [99], while studies show that the corpus callosum increases in volume with age [100]–[104].

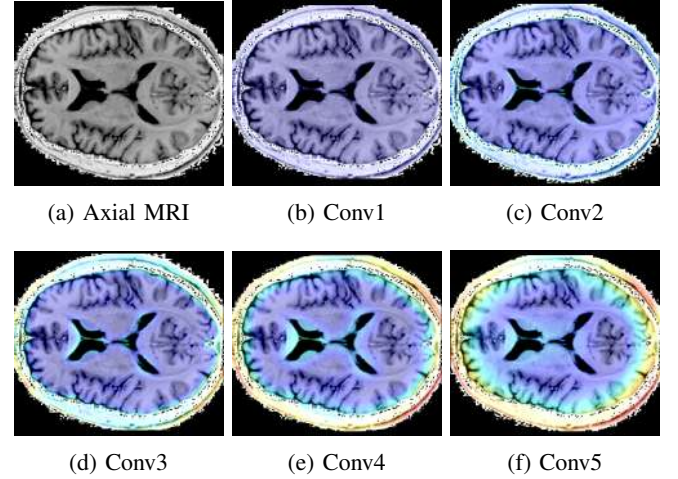


Fig. 21: Healthy brain from the IXI dataset [29] and the important features of BrainAgeNet's convolutional layers.

When utilizing the pure CNN model (the model without the MLP Age layers), BrainAgeNet can predict the person's chronological age with a mean MAE of 14.89 years and a mean RMSE of 17.96 years on the testing set. Gaussian filtering and non-local means denoising were able to reduce the errors to a MAE of 12.65 and RMSE of 15.38. Finally, incorporating the biological sex information, fusing the CNN and MLP models into one and creating the final version of the BrainAgeNet (CNN-MLP, see Fig. 12), the model predicts the chronological age with a mean MAE of 11.10 years and an RMSE of 14.20. The model manages to pick up some of the brain's features that relate to age. However the model has a proclivity to focus on the shape of the cranium instead of the actual age. This results in a model that does not focus on the actual brain, which could negatively impact reliability. More complex models are needed for this task, such as 3D Convolutions and fusions of neural networks with attention layers which will focus on the important regions of the 3D MRI image.

C. ECGNet - Arrhythmia Classification

The best and fastest convergence in the ECGNet model is achieved with the Adam optimizer using a learning rate of 0.0001. Categorical cross-entropy is used as our loss function and categorical accuracy as a performance metric. The model is trained with a batch size of 256, for 20 epochs. Ridge regularization of 0.0001 is added to the dense layers to prevent overfitting.

Figures 22, 23 display the losses and performance of the model per epoch. The model converges steadily with the choice of optimizer, learning rate and regularization we have made, with a small standard deviation in the training set. The testing set has a higher standard deviation per epoch both in its loss and its performance metric. As shown in Table II the ECGNet distinguishes between all classes of arrhythmia successfully. The model displays almost unit precision, recall and F1 for all classes. To visualize the model's performance we will again, utilize Grad-CAM. That way we will better understand the attention mechanism of the convolutional layers.

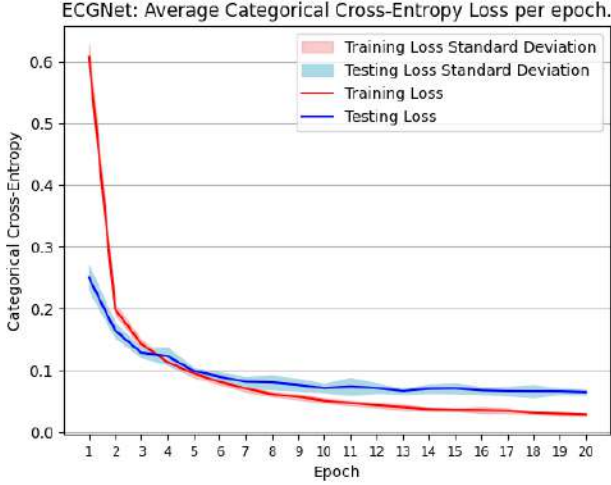


Fig. 22: Training/Testing Loss per epoch for the ECGNet.

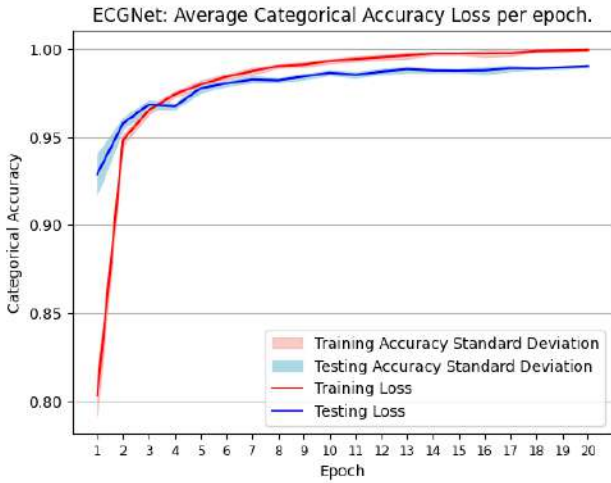


Fig. 23: Training/Testing Accuracy per epoch for the ECGNet.

1) *Normal ECG Feature Visualization*: Fig. 24 displays a normal ECG from the testing MIT-BIH [30] dataset. The intervals between each QRS complex are consistent, and all indicators are normal. The first convolutional layer focuses strictly on the R-wave (Fig. 24b). The second convolutional layer expands its focus on both the R and Q waves and the region before and after them (Fig. 24c). The third convolutional layer covers the entire QRS complex including the PR and QT intervals (Fig. 24d). Finally the fourth layer focuses on

TABLE II: ECGNet Classification Performance.

Arrhythmia Type	Precision	Recall	F1-Score
N	0.99	1.00	1.00
L	1.00	1.00	1.00
R	1.00	1.00	1.00
V	1.00	1.00	1.00
A	0.99	1.00	0.99
Accuracy 1.00 (5000 samples)			
Macro avg	1.00	1.00	1.00
Weighted avg	1.00	1.00	1.00

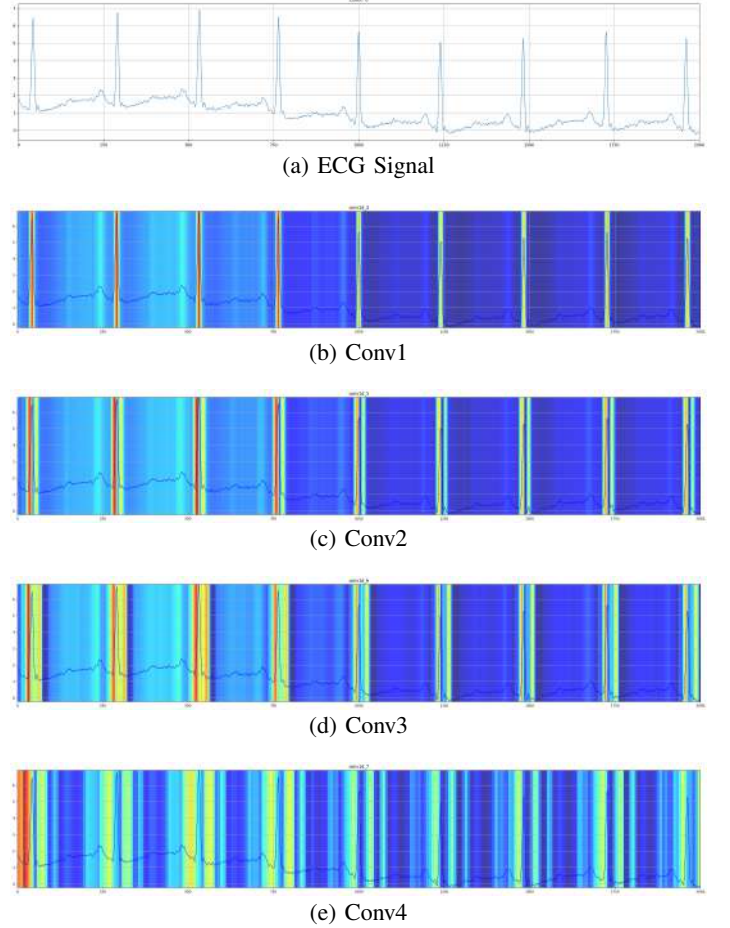


Fig. 24: Normal ECG Feature Visualization.

the regions outside the QRS complex. The results of these convolutional layers are satisfactory, since ECGNet effectively directs its attention to the QRS complex and adjacent regions, mirroring the approach the human eye would take in analyzing the same ECG segments.

2) *LBBB Feature Visualization*: Fig. 25 shows a LBBB V6 ECG. The most important characteristics of the LBBB are the deep wide rS pattern in the V1 wave with the absence of Q-waves and the RsR' pattern ('M'-wave) in V5 or V6 (Notice how the waves are in the opposite order compared to RBBB). In a normal ECG the right bundle depolarizes first and then sends its signal to the left bundle. The absence of the Q-wave

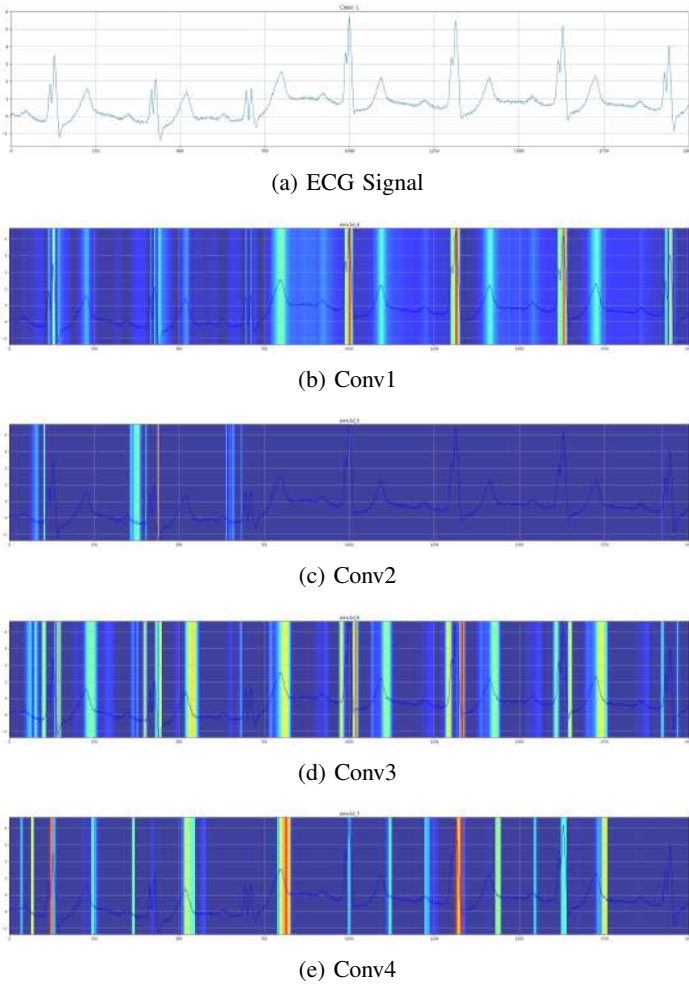


Fig. 25: LBBB V6 ECG Feature Visualization.

in each QRS complex displays the block of the left branch of the heart, since now, the right branch sends an electric signal to the left side. In the V6 ECG (Fig. 25a), the first and second convolutional layers shift their attention directly on the RsR' pattern (Fig. 25b, 25c), which is a classical presentation of LBBB. Notice the left axis deviation since the electrical current is flowing from the right branch to the left one. The third and fourth convolutional layers focus on the RsR' patterns and the absence of the Q-waves, as well as the T-waves (Fig. 25d, 25e).

Fig. 26 displays the VR-wave of an LBBB. We immediately notice the deep wide S-waves resulting from the delayed activation of the left ventricle and the abnormal depolarization spreading from the right to the left. Additionally there are absent or small R waves before the QRS complex. The convolutional layers successfully focus on the bizarre QRS complex of the VR-waves (Fig. 26b, 26c, 26d, 26e) and correctly classify it as a left bundle branch block.

3) RBBB Feature Visualization: Fig. 27 displays the V1 pulse of an RBBB irregular heartbeat. We notice a clear RsR' pattern ('M'-wave) in the QRS complexes of V1, a clear indicator of RBBB. In the V6 lateral leads we additionally

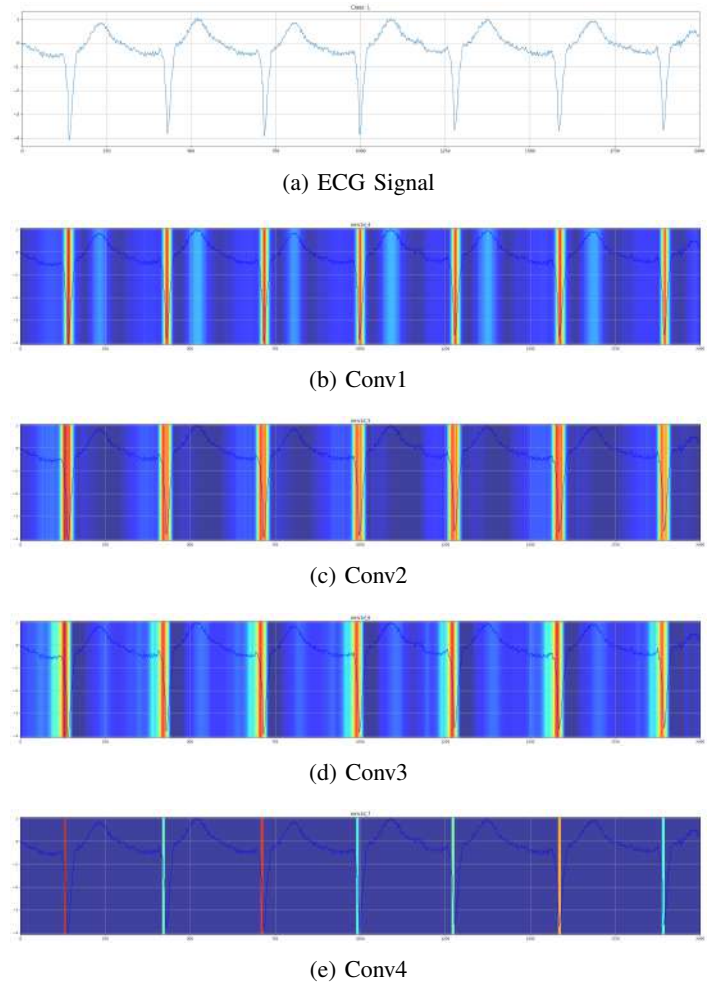


Fig. 26: LBBB VR ECG Feature Visualization.

notice a wide "slurred" S-wave (Fig. 28). In the V1 ECG the first three convolutional layers (Fig. 27b, 27c, 27d), emphasize on the RsR' pattern ('M'-wave) while in lateral V6, the first convolutional layer focuses on the R-peak (Fig. 28b), along with the wide S-wave. The second and third convolutional layers of V6 intensify their attention to the entire QRS complex, giving more attention to the wide S-wave (Fig. 28c, 28d). The fourth and final convolutional layer focuses solely on the R-peaks and the S-waves (Fig. 28e).

4) PVC Feature Visualization: Fig. 29 displays the PVC ECG feature visualization (Specifically a Bigeminal PVC - The ventricular QRS appears after every normal QRS). In Fig. 29a we notice three series of a normal P-wave followed by a normal QRS complex with each normal QRS complex followed by a wide "unnatural" QRS complex with ST suppression, T-wave inversion and a dominant R-wave. Additionally, another indicator of PVCs is the hidden P-waves of the SA-Node, since the wide ventricular rhythms suppress them in the ECG. Finally, directing our attention to 29a, we notice that the normal sinus rhythm is interrupted because of the ventricular beats and that the normal QRS complexes appear distant from each other. The first and second convolutional layers 29b, 29c focus on all QRS complexes with emphasis on the ventricular

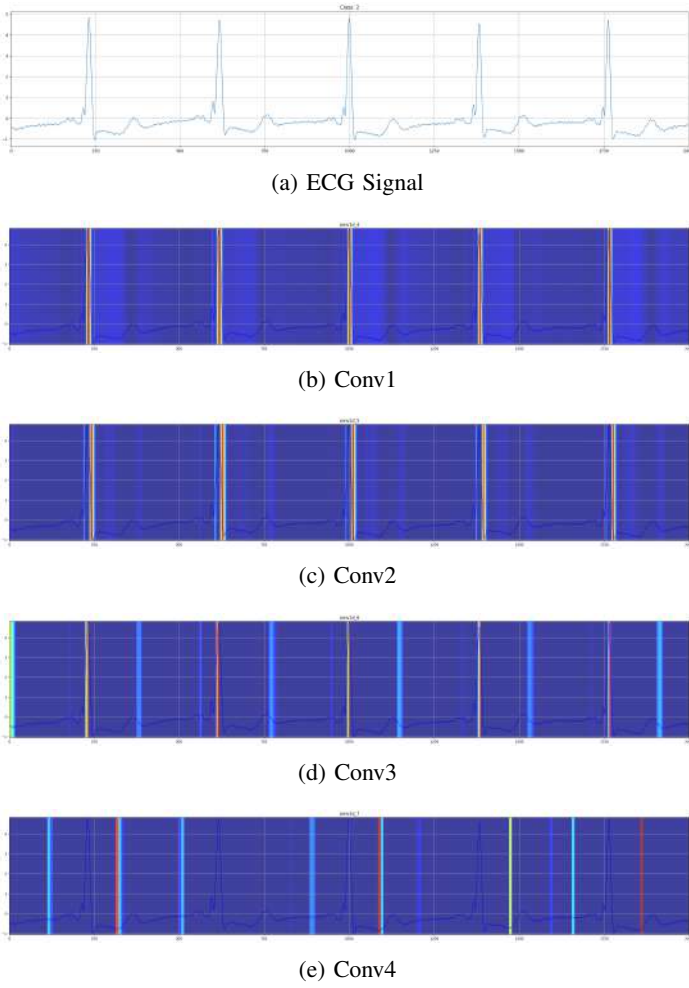


Fig. 27: RBBB V1 ECG Feature Visualization.

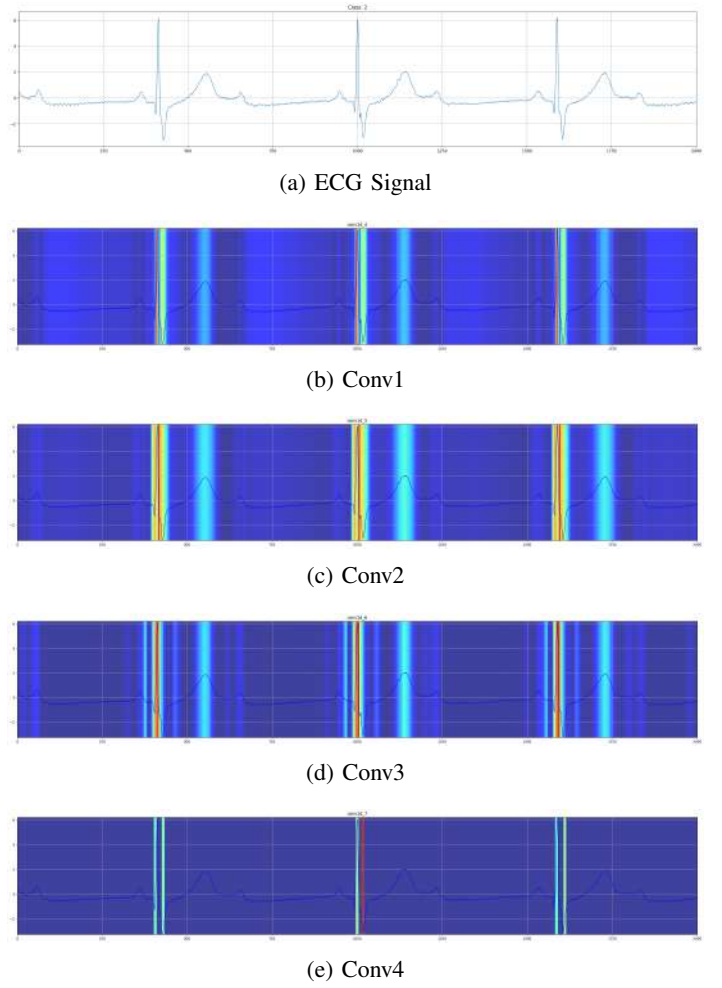


Fig. 28: RBBB V6 ECG Feature Visualization.

rhythms. Convolutional layers 3 and 4 (Fig. 29d, 29e), intensify their focus on the ventricular QRS complexes possibly noticing the absent P-waves which should follow every QRS complex. Additionally, more attention is given to the ST-suppression and T-wave inversion.

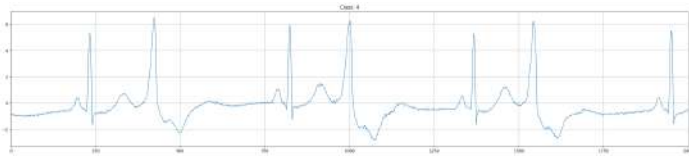
5) *PAC Feature Visualization*: The fifth and final arrhythmia type, premature atrial contraction is depicted in Fig. 30. The ectopic focus of the atria produces electrical pulses which come into conflict with electrical currents of the SA-Node. Depending on the location of the ectopic focus, the P-wave morphology before the QRS complex can appear altered or even fully inverted. Such an example is presented in Fig. 30a. Notice how some QRS complexes appear after an inverse P-wave. Additionally QRS complexes come in "pairs", with a delay in the normal sinus rhythm of the heart (compensatory pauses / inconsistent P-P intervals). This is the time the sinus node needs to reset its timing, leading to a longer interval before the next normal P-wave. Convolutional layer 1 (Fig. 30b) emphasizes the inverse P-waves leading up to the QRS complex, as well as the delay between the QRS complexes. The second convolutional layer focuses on the R-peaks with slight attention to the areas of the first layer (Fig. 30c). The third layer strongly focuses on the differences between

the normal QRS complexes and those caused by PACs (Fig. 30d). Finally the fourth layer focuses on the abnormal P-wave leading up to the PAC QRS and the absence of the QRS that should have followed (Fig. 30e).

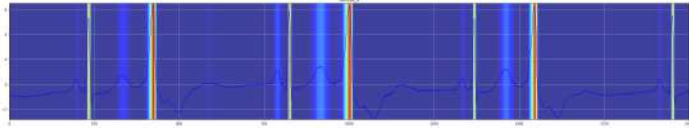
ECGNet displays outstanding performance in understanding the short-term dependencies of the ECG waves. Its attention focuses on the most important features of the wave, understanding the intricate features and structure of the data. Convolutional networks and other attention mechanisms like LSTM and RNN networks show great promise in their usage in clinical applications. Fusion of different attention mechanisms like a combination of LSTM, Bi-LSTM for long term dependencies and convolutional networks for short term dependencies within the data can be used to predict more complex and hard to detect arrhythmia types such as combinations of different arrhythmias and subcategories of arrhythmia types.

VIII. DISCUSSION - CONCLUSION

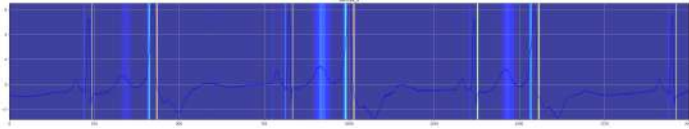
In conclusion: We propose a CNN model, the *Tumor-CNN*, that is able to classify brain MRI images into different categories of tumors, achieving an prediction accuracy of 98.78%. The model focuses successfully on the tumor regions



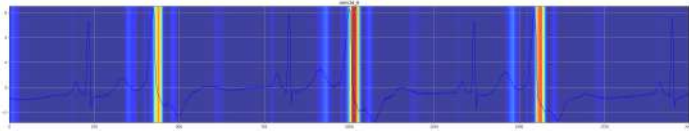
(a) ECG Signal



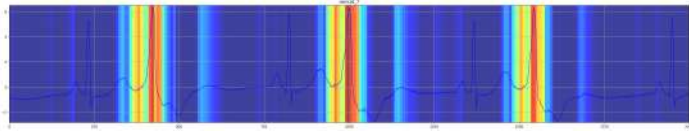
(b) Conv1



(c) Conv2

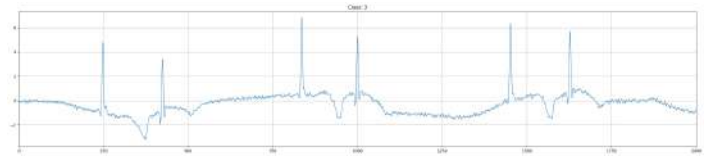


(d) Conv3

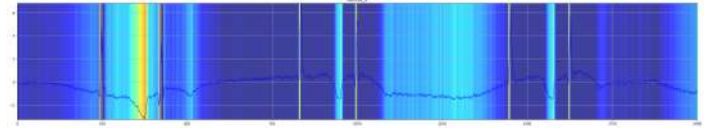


(e) Conv4

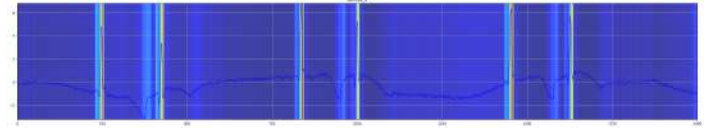
Fig. 29: PVC ECG Feature Visualization.



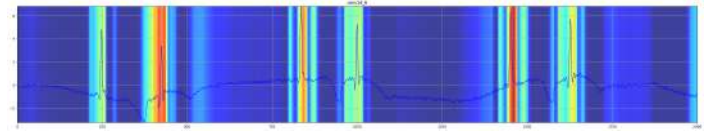
(a) ECG Signal



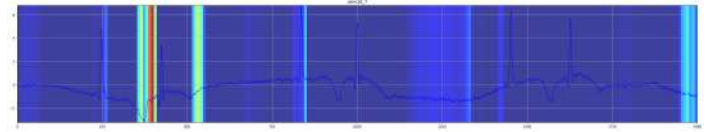
(b) Conv1



(c) Conv2



(d) Conv3



(e) Conv4

Fig. 30: PAC ECG Feature Visualization.

of the brain providing an accurate diagnosis for all types of brain tumors. The CNN-MLP *BrainAgeNet* aims to predict the human chronological age using healthy brain MRI images. The performance metrics can be improved by incorporating noise reduction/filtering algorithms and normalization techniques, and by using other predictors such as the biological sex of the patient. The model predicts the chronological age with a mean absolute error of 11.10 years and shows promise and potential in improvement with attention mechanisms which will uncover the latent structure of the brain MRI images, when predicting the human age. The one dimensional convolutional network for automatic arrhythmia classification, ECGNet, displays exceptional ability in classifying arrhythmia types and uncovering the latent structure of the data. Through visualization, which aims to increase interpretability, we confirm that the model focuses on the important parts of the ECG, in a manner similar to the one of a physician. ECGNet displays minimal error in its prediction, with an accuracy rate of 99.69%.

This study has delved into the numerous applications of neural networks in the field of medicine and aims to be a stepping stone of its improvement. The models proposed are simple, yet powerful and with potential to improve their performance through architecture experimentation, noise re-

duction, filtering algorithms and attention mechanisms that will minimize their weaknesses. Additionally the success of these models indicates a promising future for AI in enhancing the diagnostic procedure, the model interpretability and the quality of healthcare.

REFERENCES

- [1] Chao-Lung Yang, Yulius Harjoseputro, and Yung-Yao Chen. A hybrid approach of simultaneous segmentation and classification for medical image analysis. *Multimedia Tools and Applications*, 2024.
- [2] Zdravko Marinov, Paul F. Jäger, Jan Egger, Jens Kleesiek, and Rainer Stiefelhagen. Deep interactive segmentation of medical images: A systematic review and taxonomy. *arXiv preprint arXiv:2311.13964*, 2023.
- [3] Zhi Sheng, Kevin Pridham, Kasen Hutchings, and Patrick Beck. Scientists uncover promising treatment target for resistant brain cancer. *ScienceDaily*, 2024.
- [4] Timothy Cloughesy, Robert M. Prins, Patrick Wen, and et al. Neoadjuvant immunotherapy in recurrent glioblastoma patients. *Nature Medicine*, 30(3):123–130, 2024.
- [5] Yujun Hou, Xiuli Dan, Mansi Babbar, Yong Wei, Steen G Hasselbalch, Deborah L Croteau, and Vilhelm A Bohr. Ageing as a risk factor for neurodegenerative disease. *Nat Rev Neurol*, 15(10):565–581, 2019.
- [6] Rachael I Scahill, Chris Frost, Rhian Jenkins, Jennifer L Whitwell, Martin N Rossor, and Nick C Fox. A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Archives of Neurology*, 60(7):989–994, 2003.
- [7] Anders M. Fjell, Lars T. Westlye, Håkon Grydeland, and et al. Longitudinal changes in cortical thickness, surface area, and volume across the adult life span. *Journal of Neuroscience*, 2014.
- [8] Carlos López-Otín, Maria A Blasco, Linda Partridge, Manuel Serrano, and Guido Kroemer. The hallmarks of aging. *Cell*, 153(6):1194–1217, 2013.
- [9] Nikolas Rose. The human sciences in a biological age. *Theory, Culture & Society*, 30(1):3–34, 2013.
- [10] C. E. Cofey et al. Sex differences in brain aging: A quantitative magnetic resonance imaging study. *Arch. Neurol.*, 55:169–179, 1998.
- [11] K. P. Cosgrove, C. M. Mazure, and J. K. Staley. Evolving knowledge of sex differences in brain structure, function, and chemistry. *Biol. Psychiatry*, 62:847–855, 2007.
- [12] J. Xu et al. Gender effects on age-related changes in brain structure. *Am. J. Neuroradiol.*, 21:112–118, 2000.
- [13] J. B. Becker et al. Strategies and methods for research on sex differences in brain and behavior. *Endocrinology*, 146:1650–1673, 2005.
- [14] E. Vegeto et al. The role of sex and sex hormones in neurodegenerative diseases. *Endocr. Rev.*, 41:273–319, 2020.
- [15] A. M. Fjell et al. Minute effects of sex on the aging brain: A multisample magnetic resonance imaging study of healthy aging and alzheimer’s disease. *J. Neurosci.*, 29:8774–8783, 2009.
- [16] James H Cole and Katja Franke. Predicting age using neuroimaging: Innovative brain ageing biomarkers. *Trends in Neuroscience*, 40(12):681–690, 2017.
- [17] Katja Franke, Eileen Luders, Arne May, Marko Wilke, and Christian Gaser. Brain maturation: predicting individual brainage in children and adolescents using structural mri. *Neuroimage*, 63:1305–1312, 2012.
- [18] Katja Franke, Gabriel Ziegler, Stefan Kloppel, Christian Gaser, and Alzheimer’s Disease Neuroimaging Initiative. Estimating the age of healthy subjects from t1-weighted mri scans using kernel methods: exploring the influence of various parameters. *Neuroimage*, 50:883–892, 2010.
- [19] Andreas B. Storsve and et al. Differential longitudinal changes in cortical thickness, surface area, and volume across the adult life span: Regions of accelerating and decelerating change. *Journal of Neuroscience*, 34:8488–8498, 2014.
- [20] Elizabeth R. Sowell, Paul M. Thompson, and Arthur W. Toga. Mapping changes in the human cortex throughout the span of life. *Neuroscientist*, 10:372–392, 2004.
- [21] Maria Di Cesare, Pablo Perel, Sarah Taylor, Chodziwadziwa Kabudula, Honor Bixby, Thomas A. Gaziano, David V. McGhie, John Mwangi, Boris Pervan, Jagat Narula, Daniel Pineiro, and Fausto J. Pinto. The heart of the world. *Global Heart*, 19(1):11, Jan 2024.
- [22] Heart Rhythm Society. Understanding heart rhythm disorders. *Up-Beat.org*, 2024. Accessed on 31 May 2024.
- [23] Zhao He and et al. Real-time mri-guided brain intervention using lsfp-net. *Nature Communications*, 14:8257, 2023.
- [24] Zhongyu Li. D-Imbmap: a fully automated deep-learning pipeline for whole-brain profiling of neural circuitry. *Nature Methods*, 20:1593–1604, 2023.
- [25] Edward Choi, Andy Schuetz, Walter F Stewart, and Jimeng Sun. Using recurrent neural network models for early detection of heart failure onset. *Journal of the American Medical Informatics Association*, 24(2):361–370, 08 2016.
- [26] Kang Wang, Adrija Mamidipalli, Tara Retson, Naeim Bahrami, Kyle Hasenstab, Kevin Blansit, Emily Bass, Timoteo Delgado, Guilherme Cunha, Michael S. Middleton, Rohit Loomba, Brent A. Neuschwander-Tetri, Claude B. Sirlin, and Albert and Hsiao. Automated ct and mri liver segmentation and biometry using a generalized convolutional neural network. *Radiology: Artificial Intelligence*, 1(2):180022, 2019.
- [27] Yanrui Jin, Zhiyuan Li, Mengxiao Wang, Jinlei Liu, Yuanyuan Tian, Yunqing Liu, Xiaoyang Wei, Liqun Zhao, and Chengliang Liu. Cardiologist-level interpretable knowledge-fused deep neural network for automatic arrhythmia diagnosis. *Communications Medicine*, 4(31), 2024.
- [28] Jyotismita Chaki and Marcin Wozniak. Brain tumor mri dataset. 2023.
- [29] Ixi dataset. <https://brain-development.org/ixi-dataset/>, Accessed: 2024. Information eXtraction from Images (EPSRC GR/S21533/02).
- [30] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, ..., and H. E. Stanley. Physiobank, physiokit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.
- [31] M. Ahlstrom and W. Tompkins. Digital filters for real-time ecg signal processing using microprocessors. *IEEE Transactions on Biomedical Engineering*, 32(9):708–713, 1985.
- [32] B. N. Singh and A. K. Tiwari. Optimal selection of wavelet basis function applied to ecg signal denoising. *Digital Signal Processing*, 16(3):275–287, 2006.
- [33] Quinn T Ostrom, Haley Gittleman, Gabrielle Truitt, Alexander Boscia, Carol Kruchko, and Jill S Barnholtz-Sloan. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011–2015. *Neuro-Oncology*.
- [34] Martin J.B. Taphoorn, Eefje M. Sizoo, and Andrew Bottomley. Review on quality of life issues in patients with primary brain tumors. *The Oncologist*, 15(6):618–626, 2010.
- [35] J Wiemels, M Wrensch, and EB Claus. Epidemiology and etiology of meningioma. *J Neurooncol*, 99(3):307–14, Sep 2010.
- [36] YS Lim, MK Kim, BJ Park, TS Kim, and YJ Lim. Long term clinical outcomes of malignant meningiomas. *Brain Tumor Res Treat*, 1(2):85–90, Oct 2013. [PMC free article] [PubMed].
- [37] QT Ostrom, G Cioffi, H Gittleman, N Patil, K Waite, C Kruchko, and JS Barnholtz-Sloan. Cbtrus statistical report: Primary brain and other central nervous system tumors diagnosed in the united states in 2012-2016. *Neuro Oncol*, 21(Suppl 5):v1–v100, Nov 2019. [PMC free article] [PubMed].
- [38] Brigham and Women’s Hospital. Meningiomas: Understanding the most common primary brain tumor in adults. *Brigham and Women’s Hospital Medical Publications*, 2024. Accessed on 6th June 2024.
- [39] Cleveland Clinic Medical Professional. Understanding meningioma: Types, symptoms, and treatment. *Cleveland Clinic Health Library*, 2024. Available online: <https://my.clevelandclinic.org/health/diseases/17990-meningioma>.
- [40] WS Butt, LJ Weisbrod, and VN Srinivasan. *Pituitary Cancer*. StatPearls [Internet]. StatPearls Publishing, Treasure Island (FL), Jan 2023. Updated 2023 Aug 8. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK559099/>.
- [41] ME Molitch. Diagnosis and treatment of pituitary adenomas: A review. *JAMA*, 317(5):516–524, Feb 2017. [PubMed].
- [42] S Melmed. Pituitary-tumor endocrinopathies. *N Engl J Med*, 382(10):937–950, Mar 2020.
- [43] W Saeger, DK Lüdecke, M Buchfelder, R Fahlbusch, HJ Quabbe, and S Petersenn. Pathohistological classification of pituitary tumors: 10 years of experience with the german pituitary tumor registry. *Eur J Endocrinol*, 156(2):203–216, Feb 2007.

- [44] G Raverot, P Burman, A McCormack, A Heaney, S Petersenn, V Popovic, J Trouillas, and OM Dekkers. European society of endocrinology clinical practice guidelines for the management of aggressive pituitary tumours and carcinomas. *Eur J Endocrinol*, 178(1):G1–G24, Jan 2018. European Society of Endocrinology.
- [45] I Shimon. Metastatic spread to the pituitary. *Neuroendocrinology*, 110(9-10):805–808, 2020.
- [46] SL Taylor, JA Barakos, GR Harsh, and CB Wilson. Magnetic resonance imaging of tuberculum sellae meningiomas: preventing preoperative misdiagnosis as pituitary macroadenoma. *Neurosurgery*, 31(4):621–627, Oct 1992. discussion 627.
- [47] R Pivonello, A Colao, C Di Somma, G Faccioli, M Klain, A Faggiano, M Salvatore, and G Lombardi. Impairment of bone status in patients with central diabetes insipidus. *J Clin Endocrinol Metab*, 83(7):2275–2280, Jul 1998.
- [48] JN Domingue and CB Wilson. Pituitary abscesses. report of seven cases and review of the literature. *J Neurosurg*, 46(5):601–608, May 1977.
- [49] B Robinson. Intracellular abscess after transsphenoidal pituitary adenectomy. *Neurosurgery*, 12(6):684–686, Jun 1983.
- [50] I Ciric, M Mikhael, T Stafford, L Lawson, and R Garces. Transsphenoidal microsurgery of pituitary macroadenomas with long-term follow-up results. *J Neurosurg*, 59(3):395–401, Sep 1983.
- [51] WT Harkness and M Hicks. *Right Bundle Branch Block*. StatPearls [Internet]. StatPearls Publishing, Treasure Island (FL), Jan 2023. Updated 2023 Aug 8. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK507872/>.
- [52] D Scherbak and GJ Hicks. *Left Bundle Branch Block*. StatPearls [Internet]. StatPearls Publishing, Treasure Island (FL), Jan 2023. Updated 2023 Jul 25. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK482167/>.
- [53] P Francia, C Balla, F Paneni, and M Volpe. Left bundle-branch block—pathophysiology, prognosis, and clinical management. *Clin Cardiol*, 30(3):110–115, Mar 2007.
- [54] MH Nikoo, A Aslani, and MV Jorat. Lbbb: State-of-the-art criteria. *Int Cardiovasc Res J*, 7(2):39–40, Jun 2013.
- [55] B Surawicz, R Childers, BJ Deal, LS Gettes, JJ Bailey, A Gorgels, EW Hancock, M Josephson, P Kligfield, JA Kors, P Macfarlane, JW Mason, DM Mirvis, P Okin, O Pahlm, PM Rautaharju, G van Herpen, GS Wagner, H Wellens, American Heart Association Electrocardiography, Council on Clinical Cardiology Arrhythmias Committee, American College of Cardiology Foundation, and Heart Rhythm Society. Aha/accf/hrs recommendations for the standardization and interpretation of the electrocardiogram: part iii: intraventricular conduction disturbances: a scientific statement from the american heart association electrocardiography and arrhythmias committee, council on clinical cardiology; the american college of cardiology foundation; and the heart rhythm society. endorsed by the international society for computerized electrocardiology. *J Am Coll Cardiol*, 53(11):976–981, Mar 17 2009.
- [56] D Conen, M Adam, F Roche, JC Barthelemy, D Felber Dietrich, M Imboden, N Künzli, A von Eckardstein, S Regenass, T Hornemann, T Rochat, JM Gaspoz, N Probst-Hensch, and D Carballo. Premature atrial contractions in the general population: frequency and risk factors. *Circulation*, 126(19):2302–2308, Nov 06 2012.
- [57] J Heaton and S Yandrapalli. *Premature Atrial Contractions*. StatPearls [Internet]. StatPearls Publishing, Treasure Island (FL), Jan 2023. Updated 2023 Aug 8. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK559204/>.
- [58] BH Chong, V Pong, KF Lam, S Liu, ML Zuo, YF Lau, CP Lau, HF Tse, and CW Siu. Frequent premature atrial complexes predict new occurrence of atrial fibrillation and adverse cardiovascular events. *Europace*, 14(7):942–947, Jul 2012.
- [59] BT Huang, FY Huang, Y Peng, YB Liao, F Chen, TL Xia, XB Pu, and M Chen. Relation of premature atrial complexes with stroke and death: Systematic review and meta-analysis. *Clin Cardiol*, 40(11):962–969, Nov 2017.
- [60] W Qureshi, AJ Shah, T Salahuddin, and EZ Soliman. Long-term mortality risk in individuals with atrial or ventricular premature complexes (results from the third national health and nutrition examination survey). *Am J Cardiol*, 114(1):59–64, Jul 01 2014.
- [61] C Gianni, JD Burkhardt, C Trivedi, S Mohanty, and A Natale. The role of the purkinje network in premature ventricular complex-triggered ventricular fibrillation. *J Interv Card Electrophysiol*, 52(3):375–383, Aug 2018.
- [62] A Frigy, E Csiki, C Caraşca, IA Szabó, and VD Moga. Autonomic influences related to frequent ventricular premature beats in patients without structural heart disease. *Medicine (Baltimore)*, 97(28):e11489, Jul 2018.
- [63] K Farzam and JR Richards. *Premature Ventricular Contraction*. StatPearls [Internet]. StatPearls Publishing, Treasure Island (FL), Jan 2023. Updated 2023 Aug 8. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK532991/>.
- [64] YG Kim, YY Choi, KD Han, KJ Min, HY Choi, J Shim, JI Choi, and YH Kim. Premature ventricular contraction increases the risk of heart failure and ventricular tachyarrhythmia. *Sci Rep*, 11(1):12698, Jun 16 2021.
- [65] A Lador, JR Giudicessi, WM Barake, PA Noseworthy, S Kapa, CJ McLeod, BC Cannon, SJ Asirvatham, and MJ Ackerman. Premature ventricular contraction-triggered ventricular fibrillation and sudden cardiac arrest in the young. *JACC Clin Electrophysiol*, 8(3):380–382, Mar 2022. Epub 2021 Nov 24.
- [66] A Saurav, A Smer, A Abuzaid, O Bansal, and H Abuissa. Premature ventricular contraction-induced cardiomyopathy. *Clin Cardiol*, 38(4):251–258, Apr 2015. Epub 2015 Feb 10.
- [67] WT Hubel and TN Wiesel. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962.
- [68] T Lindeberg. Scale invariant feature transform, 2012.
- [69] N Dalal and B Triggs. Histograms of oriented gradients for human detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 886–893, 2005.
- [70] T Ahonen, A Hadid, and M Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, 2006.
- [71] Ergün Yücesoy. Speaker age and gender recognition using 1d and 2d convolutional neural networks. *Neural Computing and Applications*, 36:3065–3075, 2024.
- [72] D. Gilbert Chandra, U. Srinivasulu Reddy, G. Uma, et al. Group normalization-based 2d-convolutional neural network for intelligent bearing fault diagnosis. *J Braz. Soc. Mech. Sci. Eng.*, 45(584), 2023.
- [73] A. Pandey, P. Kumar, and S. Prasad. 2d convolutional lstm-based approach for human action recognition on various sensor data. In V. Bhatija, XS. Yang, J. Chun-Wei Lin, and R. Das, editors, *Intelligent Data Engineering and Analytics*, volume 327 of *Smart Innovation, Systems and Technologies*. Springer, Singapore, 2023.
- [74] A. Joshi, R. Golchha, R.N. Giri, R.R. Janghel, H. Govil, and S.K. Pandey. A hybrid approach using wavelet and 2d convolutional neural network for hyperspectral image classification. In *Machine Intelligence Techniques for Data Analysis and Signal Processing*, volume 997 of *Lecture Notes in Electrical Engineering*. Springer, Singapore, 2023.
- [75] Arnaud AA Setio, Francesco Ciompi, Geert Litjens, et al. Pulmonary nodule detection in ct images: false positive reduction using multi-view convolutional networks. *IEEE Transactions on Medical Imaging*, 35:1160–1169, 2016.
- [76] Koichiro Yasaka, Hirofumi Akai, Osamu Abe, and Shigehiro Kiryu. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced ct: a preliminary study. *Radiology*, 286:887–896, 2018.
- [77] S Kiranyaz, T Ince, R Hamila, and M Gabbouj. Convolutional neural networks for patient-specific ecg classification. In *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, 2015.
- [78] S Kiranyaz, T Ince, and M Gabbouj. Real-time patient-specific ecg classification by 1-d convolutional neural networks. *IEEE Trans. Biomed. Eng.*, 63:664–675, 2016.
- [79] S Kiranyaz, T Ince, and M Gabbouj. Personalized monitoring and advance warning system for cardiac arrhythmias. *Sci. Rep.*, 7, 2017.
- [80] O Avci, O Abdeljaber, S Kiranyaz, M Hussein, and DJ Inman. Wireless and real-time structural damage detection: a novel decentralized method for wireless sensor networks. *J. Sound Vib.*, 2018.

- [81] O Avci, O Abdeljaber, S Kiranyaz, and D Inman. Structural damage detection in real time: implementation of 1d convolutional neural networks for shm applications. In C Niezrecki, editor, *Struct. Heal. Monit. Damage Detect. Vol. 7 Proc. 35th IMAC, A Conf. Expo. Struct. Dyn. 2017*, pages 49–54. Springer International Publishing, Cham, 2017.
- [82] O Abdeljaber, O Avci, S Kiranyaz, M Gabbouj, and DJ Inman. Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. *J. Sound Vib.*, 388, 2017.
- [83] O Avci, O Abdeljaber, S Kiranyaz, B Boashash, H Sodano, and DJ Inman. Efficiency validation of one dimensional convolutional neural networks for structural damage detection using a shm benchmark data, 2018.
- [84] O Abdeljaber, O Avci, MS Kiranyaz, B Boashash, H Sodano, and DJ Inman. 1-d cnns for structural damage detection: verification on a structural health monitoring benchmark data. *Neurocomputing*, 2017.
- [85] T Ince, S Kiranyaz, L Eren, M Askar, and M Gabbouj. Real-time motor fault detection by 1-d convolutional neural networks. *IEEE Trans. Ind. Electron.*, 63:7067–7075, 2016.
- [86] S Kiranyaz, A Gastli, L Ben-Brahim, N Alemadi, and M Gabbouj. Real-time fault detection and identification for mmc using 1d convolutional neural networks. *IEEE Trans. Ind. Electron.*, 2018.
- [87] O Abdeljaber, S Sassi, O Avci, S Kiranyaz, I Abulrahman, and M Gabbouj. Fault detection and severity identification of ball bearings by online condition monitoring. *IEEE Trans. Ind. Electron.*, 2018.
- [88] L Eren, T Ince, and S Kiranyaz. A generic intelligent bearing fault diagnosis system using compact adaptive 1d cnn classifier. *J. Signal Process. Syst.*, 91:179–189, 2019.
- [89] L Eren. Bearing fault detection by one-dimensional convolutional neural networks. *Math. Probl. Eng.*, 2017.
- [90] N Qian. On the momentum term in gradient descent learning algorithms. *Neural Networks*, 12:145–151, 1999.
- [91] J Duchi, E Hazan, and Y Singer. Adaptive subgradient methods for online learning and stochastic optimization. In *COLT 2010-23rd Conf. Learn. Theory*, pages 257–269, 2010.
- [92] T Tieleman and G Hinton. Lecture 6.5 - rmsprop, neural networks for machine learning — coursera. n.d.
- [93] Diederik P Kingma and Jimmy Lei Ba. Adam: a method for stochastic optimization. In *AIP Conf. Proc.*, volume 1631, pages 58–62, 2014.
- [94] S Ruder. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.
- [95] P. Cortez and M. J. Embrechts. Using sensitivity analysis and visualization techniques to open black box data mining models. *Inf. Sci.*, 225:1–17, 2013.
- [96] R. R. Selvaraju et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 618–626, 2017.
- [97] J. R. Nesselroade, S. M. Stigler, and P. B. Baltes. Regression toward the mean and the study of change. *Psychol. Bull.*, 88:622, 1980.
- [98] J. Wang et al. Gray matter age prediction as a biomarker for risk of dementia. *Proc. Natl. Acad. Sci. U. S. A.*, 116:21213–21218, 2019.
- [99] Anne Hafkemeijer, Irmhild Altmann-Schneider, Anton JM de Craen, P Eline Slagboom, Jeroen van der Grond, and Serge ARB Rombouts. Associations between age and gray matter volume in anatomical brain networks in middle-aged to older adults. *Aging Cell*, 13(6):1068–1074, 2014.
- [100] M. M. Tanaka-Arakawa, M. Matsui, C. Tanaka, A. Uematsu, S. Uda, K. Miura, T. Sakai, and K. Noguchi. Developmental changes in the corpus callosum from infancy to early adulthood: a structural magnetic resonance imaging study. *PLoS One*, 10(3):e0118760, Mar 2015. Erratum in: *PLoS One*. 2015 Jul 13;10(7):e0133090. doi: 10.1371/journal.pone.0133090.
- [101] J. N. Giedd, J. M. Rumsey, F. X. Castellanos, J. C. Rajapakse, D. Kaysen, and J. L. Rapoport. A quantitative mri study of the corpus callosum in children and adolescents. *Develop Brain Res*, 91:274–280, 1996.
- [102] A. J. Barkovich and D. Norman. Anomalies of the corpus callosum: correlation with further anomalies of the brain. *Am J Roentgenol*, 151:171–179, 1988.
- [103] J. Pujol, P. Vendrell, C. Junqué, J. L. Martí-Vilalta, and A. Capdevila. When does human brain development end? evidence of corpus callosum growth up to adulthood. *Ann Neurol*, 34:71–75, 1993.
- [104] P. Rakic and P. I. Yakovlev. Development of the corpus callosum and cavum septi in man. *J Comp Neurol*, 132:45–72, 1968.