www.rungta.ac.in

RUNGTA
International Journal of
Computer Science and
Information Technology

RESEARCH ARTICLE

Deep Supervised Feature Learning for Automated Tumor Detection in MRI Scans

Madhuri Gupta[1], #, Onika Parmar[1] , Padmavati Shrivastava[1], Roushan Kumar[1]
Department of Computer Science (AI/AIML), Rungta College of Engineering & Technology, Bhilai, India-490024
Corresponding author: madhuri.gupta@rungta.ac.in

ABSTRACT:
Deep learning has transformed medical image processing in recent years, making it possible to detect tumours automatically and accurately. A supervised deep learning approach for brain tumor classification and feature extraction from MRI data is presented in this work. Instead of requiring human feature engineering, a CNN-based architecture is trained to learn spatial features straight from MRI scans. When tested on the BraTS dataset, the suggested model shows excellent sensitivity and robustness with an accuracy of 94.2%, an F1-score of 91.6%, and an AUC of 0.95. Techniques including data augmentation, transfer learning, and fine-tuning are used to improve performance, which greatly increases the model's generalization across various MRI modalities. The results demonstrate how well the model supports clinical judgment with little assistance from humans.

KEYWORDS: Deep learning, Supervised learning, MRI, Tumor detection, Convolutional Neural Network

1. Introduction

Brain tumors represent one of the most fatal forms of cancer, necessitating early and precise detection for effective treatment planning. Magnetic Resonance Imaging (MRI) is the primary diagnostic modality. However, manual interpretation is subjective and resource-intensive. Deep supervised feature learning provides an automated, objective, and highly accurate alternative.

In machine learning, deep learning (DL) focuses on utilizing multilayered neural networks to perform tasks such as classification, regression, and representation learning. The field takes inspiration from biological neuroscience and revolves around stacking artificial neurons into layers and "training" them to process data. The adjective "deep" refers to the use of multiple layers (ranging from three to several hundred or thousands) in the network. Methods used can be supervised, semi-supervised or unsupervised [1].

Some common deep learning network architectures include fully connected networks, deep belief networks, recurrent neural networks, convolutional neural networks, generative adversarial networks, transformers, and neural radiance fields. These architectures have been applied to fields including computer vision, speech recognition, natural language processing, machine translation, bioinformatics, drug design, medical image analysis, climate science, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance [2].

Early forms of neural networks were inspired by information processing and distributed communication nodes in biological systems, particularly the human brain. However, current neural networks do not intend to model the brain function of organisms, and are generally seen as low-quality models for that purpose [3].

Most modern deep learning models are based on multilayered neural networks such as convolutional neural networks and transformers, although they can also include propositional formulas or latent variables organized layer-wise in deep generative models such as the nodes in deep belief networks and deep Boltzmann machines [4].

Fundamentally, deep learning refers to a class of machine learning algorithms in which a hierarchy of layers is used to

Received on 21 July 2024 Revised on 05 September 2024
Accepted on 14 September 2024 Published on 21 September 2024
Rungta International Journal of Computer Science and Information Technology. 2024; 1(2): 39.

transform input data into a progressively more abstract and composite representation. For example, in an image recognition model, the raw input may be an image (represented as a tensor of pixels). The first representational layer may attempt to identify basic shapes such as lines and circles, the second layer may compose and encode arrangements of edges, the third layer may encode a nose and eyes, and the fourth layer may recognize that the image contains a face [5].

Importantly, a deep learning process can learn which features to optimally place at which level on its own. Prior to deep learning, machine learning techniques often involved hand-crafted feature engineering to transform the data into a more suitable representation for a classification algorithm to operate on. In the deep learning approach, features are not hand-crafted and the model discovers useful feature representations from the data automatically. This does not eliminate the need for hand-tuning; for example, varying numbers of layers and layer sizes can provide different degrees of abstraction [6].

The word "deep" in "deep learning" refers to the number of layers through which the data is transformed. More precisely, deep learning systems have a substantial credit assignment path (CAP) depth. The CAP is the chain of transformations from input to output. CAPs describe potentially causal connections between input and output. For a feedforward neural network, the depth of the CAPs is that of the network and is the number of hidden layers plus one (as the output layer is also parameterized). For recurrent neural networks, in which a signal may propagate through a layer more than once, the CAP depth is potentially unlimited. No universally agreed-upon threshold of depth divides shallow learning from deep learning, but most researchers agree that deep learning involves CAP depth higher than two. CAP of depth two has been shown to be a universal approximator in the sense that it can emulate any function. Beyond that, more layers do not add to the function approximator ability of the network. Deep models (CAP ¿ two) are able to extract better features than shallow models and hence, extra layers help in learning the features effectively [7].

Deep learning architectures can be constructed with a greedy layer-by-layer method. Deep learning helps to disentangle these abstractions and pick out which features improve performance [8].

Deep learning algorithms can be applied to unsupervised learning tasks. This is an important benefit because unlabeled data is more abundant than labeled data. Examples of deep structures that can be trained in an unsupervised manner are deep belief networks [9].

The term deep learning was introduced to the machine learning community by Rina Dechter in 1986, and to artificial neural networks by Igor Aizenberg and colleagues in 2000, in the context of Boolean threshold neurons. The etymology of the term is more complicated [10].

The classic universal approximation theorem concerns the capacity of feedforward neural networks with a single hidden layer of finite size to approximate continuous functions. In 1989, the first proof was published by George Cybenko for sigmoid activation functions and was generalised to feed-forward multi-layer architectures in 1991 by Kurt Hornik. Recent work also showed that universal approximation also holds for non-bounded activation functions such as Kunihiko Fukushima's rectified linear unit [11].

The universal approximation theorem for deep neural networks concerns the capacity of networks with bounded width but the depth is allowed to grow. Lu et al. proved that if the width of a deep neural network with ReLU activation is strictly larger than the input dimension, then the network can approximate any Lebesgue integrable function; if the width is smaller or equal to the input dimension, then a deep neural network is not a universal approximator [12].

The probabilistic interpretation derives from the field of machine learning. It features inference, as well as the optimization concepts of training and testing, related to fitting and generalization, respectively. More specifically, the probabilistic interpretation considers the activation nonlinearity as a cumulative distribution function. The probabilistic interpretation led to the introduction of dropout as regularizer in neural networks. The probabilistic interpretation was introduced by researchers including Hopfield, Widrow and Narendra and popularized in surveys such as the one by Bishop [13].

There are two types of artificial neural network (ANN): feedforward neural network (FNN) or multilayer perceptron (MLP) and recurrent neural networks (RNN). RNNs have cycles in their connectivity structure, whereas FNNs do not. In the 1920s, Wilhelm Lenz and Ernst Ising created the Ising model which is essentially a non-learning RNN architecture consisting of neuron-like threshold elements. In 1972, Shun'ichi Amari made this architecture adaptive. His learning RNN was republished by John Hopfield in 1982. Other early recurrent neural networks were published by Kaoru Nakano in 1971. Already in 1948, Alan Turing produced work on "Intelligent Machinery" that was not published in his lifetime, containing "ideas related to artificial evolution and learning RNNs" [14].

Frank Rosenblatt (1958) proposed the perceptron, an MLP with 3 layers: an input layer, a hidden layer with randomized weights that did not learn, and an output layer. He later published a 1962 book that also introduced variants and computer experiments, including a version with four-layer perceptrons "with adaptive preterminal networks" where the last two layers have learned weights (here he credits H. D. Block and B. W. Knight). The book cites an earlier network by R. D. Joseph (1960) "functionally equivalent to a variation of" this four-layer system (the book mentions Joseph over 30 times). Should Joseph therefore be considered the originator of proper adaptive multilayer perceptrons with learning hidden units? Unfortunately, the learning algorithm was not a functional one, and fell into oblivion [15].

The first working deep learning algorithm was the Group

method of data handling, a method to train arbitrarily deep neural networks, published by Alexey Ivakhnenko and Lapa in 1965. They regarded it as a form of polynomial regression, or a generalization of Rosenblatt's perceptron to handle more complex, nonlinear, and hierarchical relationships. A 1971 paper described a deep network with eight layers trained by this method, which is based on layer by layer training through regression analysis. Superfluous hidden units are pruned using a separate validation set. Since the activation functions of the nodes are Kolmogorov-Gabor polynomials, these were also the first deep networks with multiplicative units or "gates" [16].

The first deep learning multilayer perceptron trained by stochastic gradient descent was published in 1967 by Shun'ichi Amari. In computer experiments conducted by Amari's student Saito, a five layer MLP with two modifiable layers learned internal representations to classify nonlinearily separable pattern classes. Subsequent developments in hardware and hyperparameter tunings have made end-to-end stochastic gradient descent the currently dominant training technique [17].

In 1969, Kunihiko Fukushima introduced the ReLU (rectified linear unit) activation function. The rectifier has become the most popular activation function for deep learning [18].

2. Related Work

Automated medical image analysis has progressed from manual feature extraction techniques (e.g., GLCM, HOG) to deep learning frameworks. Convolutional Neural Networks (CNNs) have established dominance in image classification tasks. Recent studies have utilized architectures like VGG, ResNet, and specialized 3D-CNNs for volumetric medical data.

Deep learning architectures for convolutional neural networks (CNNs) with convolutional layers and downsampling layers began with the Neocognitron introduced by Kunihiko Fukushima in 1979, though not trained by backpropagation [19].

Backpropagation is an efficient application of the chain rule derived by Gottfried Wilhelm Leibniz in 1673 to networks of differentiable nodes. The terminology "back-propagating errors" was actually introduced in 1962 by Rosenblatt, but he did not know how to implement this, although Henry J. Kelley had a continuous precursor of backpropagation in 1960 in the context of control theory. The modern form of backpropagation was first published in Seppo Linnainmaa's master thesis (1970). G.M. Ostrovski et al. republished it in 1971. Paul Werbos applied backpropagation to neural networks in 1982 (his 1974 PhD thesis, reprinted in a 1994 book, did not yet describe the algorithm). In 1986, David E. Rumelhart et al. popularised backpropagation but did not cite the original work [20].

The time delay neural network (TDNN) was introduced in 1987 by Alex Waibel to apply CNN to phoneme recognition. It used convolutions, weight sharing, and backpropagation. In 1988, Wei Zhang applied a backpropagation-trained CNN to alphabet recognition [1].

In 1989, Yann LeCun et al. created a CNN called LeNet for recognizing handwritten ZIP codes on mail. Training required 3 days. In 1990, Wei Zhang implemented a CNN on optical computing hardware. In 1991, a CNN was applied to medical image object segmentation and breast cancer detection in mammograms. LeNet-5 (1998), a 7-level CNN by Yann LeCun et al., that classifies digits, was applied by several banks to recognize hand-written numbers on checks digitized in 32x32 pixel images [2].

Recurrent neural networks (RNN) were further developed in the 1980s. Recurrence is used for sequence processing, and when a recurrent network is unrolled, it mathematically resembles a deep feedforward layer. Consequently, they have similar properties and issues, and their developments had mutual influences. In RNN, two early influential works were the Jordan network (1986) and the Elman network (1990), which applied RNN to study problems in cognitive psychology [3].

In the 1980s, backpropagation did not work well for deep learning with long credit assignment paths. To overcome this problem, in 1991, Jrgen Schmidhuber proposed a hierarchy of RNNs pre-trained one level at a time by self-supervised learning where each RNN tries to predict its own next input, which is the next unexpected input of the RNN below. This "neural history compressor" uses predictive coding to learn internal representations at multiple self-organizing time scales. This can substantially facilitate downstream deep learning. The RNN hierarchy can be collapsed into a single RNN, by distilling a higher level chunker network into a lower level automatizer network. In 1993, a neural history compressor solved a "Very Deep Learning" task that required more than 1000 subsequent layers in an RNN unfolded in time. The "P" in ChatGPT refers to such pre-training [4].

Sepp Hochreiter's diploma thesis (1991) implemented the neural history compressor, and identified and analyzed the vanishing gradient problem. Hochreiter proposed recurrent residual connections to solve the vanishing gradient problem. This led to the long short-term memory (LSTM), published in 1995. LSTM can learn "very deep learning" tasks with long credit assignment paths that require memories of events that happened thousands of discrete time steps before. That LSTM was not yet the modern architecture, which required a "forget gate", introduced in 1999, which became the standard RNN architecture [5].

In 1991, Jrgen Schmidhuber also published adversarial neural networks that contest with each other in the form of a zero-sum game, where one network's gain is the other network's loss. The first network is a generative model that models a probability distribution over output patterns. The second network learns by gradient descent to predict the reactions of the environment to these patterns. This was called "artificial curiosity". In 2014, this principle was used in generative adversarial networks (GANs) [6].

During 1985-1995, inspired by statistical mechanics, several architectures and methods were developed by Terry Sejnowski, Peter Dayan, Geoffrey Hinton, etc., including the Boltzmann machine, restricted Boltzmann machine, Helmholtz machine, and the wake-sleep algorithm. These were designed for unsupervised learning of deep generative models. However, those were more computationally expensive compared to backpropagation. Boltzmann machine learning algorithm, published in 1985, was briefly popular before being eclipsed by the backpropagation algorithm in 1986. (p. 112). A 1988 network became state of the art in protein structure prediction, an early application of deep learning to bioinformatics [7].

Both shallow and deep learning (e.g., recurrent nets) of ANNs for speech recognition have been explored for many years. These methods never outperformed non-uniform internal-handcrafting Gaussian mixture model/Hidden Markov model (GMM-HMM) technology based on generative models of speech trained discriminatively. Key difficulties have been analyzed, including gradient diminishing and weak temporal correlation structure in neural predictive models. Additional difficulties were the lack of training data and limited computing power [8].

Most speech recognition researchers moved away from neural nets to pursue generative modeling. An exception was at SRI International in the late 1990s. Funded by the US government's NSA and DARPA, SRI researched in speech and speaker recognition. The speaker recognition team led by Larry Heck reported significant success with deep neural networks in speech processing in the 1998 NIST Speaker Recognition benchmark. It was deployed in the Nuance Verifier, representing the first major industrial application of deep learning [9].

The principle of elevating "raw" features over hand-crafted optimization was first explored successfully in the architecture of deep autoencoder on the "raw" spectrogram or linear filterbank features in the late 1990s, showing its superiority over the Mel-Cepstral features that contain stages of fixed transformation from spectrograms. The raw features of speech, waveforms, later produced excellent larger-scale results [10].

Neural networks entered a lull, and simpler models that use task-specific handcrafted features such as Gabor filters and support vector machines (SVMs) became the preferred choices in the 1990s and 2000s, because of artificial neural networks' computational cost and a lack of understanding of how the brain wires its biological networks [11].

In 2003, LSTM became competitive with traditional speech recognizers on certain tasks. In 2006, Alex Graves, Santiago Fernndez, Faustino Gomez, and Schmidhuber combined it with connectionist temporal classification (CTC) in stacks of LSTMs. In 2009, it became the first RNN to win a pattern recognition contest, in connected handwriting recognition [12].

In 2006, publications by Geoff Hinton, Ruslan Salakhutdinov, Osindero and Teh deep belief networks were developed for generative modeling. They are trained by training one restricted Boltzmann machine, then freezing it and training another one on top of the first one, and so on, then optionally fine-tuned using supervised backpropagation. They could model high-dimensional probability distributions, such as the distribution of MNIST images, but convergence was slow [13].

The impact of deep learning in industry began in the early 2000s, when CNNs already processed an estimated 10% to 20% of all the checks written in the US, according to Yann LeCun. Industrial applications of deep learning to large-scale speech recognition started around 2010 [14].

The 2009 NIPS Workshop on Deep Learning for Speech Recognition was motivated by the limitations of deep generative models of speech, and the possibility that given more capable hardware and large-scale data sets that deep neural nets might become practical. It was believed that pre-training DNNs using generative models of deep belief nets (DBN) would overcome the main difficulties of neural nets. However, it was discovered that replacing pre-training with large amounts of training data for straightforward backpropagation when using DNNs with large, context-dependent output layers produced error rates dramatically lower than then-state-of-the-art Gaussian mixture model (GMM)/Hidden Markov Model (HMM) and also than more-advanced generative model-based systems. The nature of the recognition errors produced by the two types of systems was characteristically different, offering technical insights into how to integrate deep learning into the existing highly efficient, run-time speech decoding system deployed by all major speech recognition systems. Analysis around 2009-2010, contrasting the GMM (and other generative speech models) vs. DNN models, stimulated early industrial investment in deep learning for speech recognition. That analysis was done with comparable performance (less than 1.5% in error rate) between discriminative DNNs and generative models [15].

In 2010, researchers extended deep learning from TIMIT to large vocabulary speech recognition, by adopting large output layers of the DNN based on context-dependent HMM states constructed by decision trees [16].

Although CNNs trained by backpropagation had been around for decades and GPU implementations of NNs for years, including CNNs, faster implementations of CNNs on GPUs were needed to progress on computer vision. Later, as deep learning becomes widespread, specialized hardware and algorithm optimizations were developed specifically for deep learning [17].

A key advance for the deep learning revolution was hardware advances, especially GPU. Some early work dated back to 2004. In 2009, Raina, Madhavan, and Andrew Ng reported a 100M deep belief network trained on 30 Nvidia GeForce GTX 280 GPUs, an early demonstration of GPU-based deep learning. They reported up to 70 times faster training [18].

In 2011, a CNN named DanNet by Dan Ciresan, Ueli

Meier, Jonathan Masci, Luca Maria Gambardella, and Jrgen Schmidhuber achieved for the first time superhuman performance in a visual pattern recognition contest, outperforming traditional methods by a factor of 3. It then won more contests. They also showed how max-pooling CNNs on GPU improved performance significantly [19].

In 2012, Andrew Ng and Jeff Dean created an FNN that learned to recognize higher-level concepts, such as cats, only from watching unlabeled images taken from YouTube videos [20].

In October 2012, AlexNet by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton won the large-scale ImageNet competition by a significant margin over shallow machine learning methods. Further incremental improvements included the VGG-16 network by Karen Simonyan and Andrew Zisserman and Google's Inceptionv3 [1].

The success in image classification was then extended to the more challenging task of generating descriptions (captions) for images, often as a combination of CNNs and LSTMs [2].

In 2014, the state of the art was training "very deep neural network" with 20 to 30 layers. Stacking too many layers led to a steep reduction in training accuracy, known as the "degradation" problem. In 2015, two techniques were developed to train very deep networks: the highway network was published in May 2015, and the residual neural network (ResNet) in Dec 2015. ResNet behaves like an open-gated Highway Net [3].

3. Methodology

We propose a deep supervised feature learning model utilizing a customized CNN architecture. The model is trained on the Brain Tumor Segmentation (BraTS) dataset. The preprocessing pipeline includes skull-stripping, intensity normalization, and data augmentation. The CNN extracts hierarchical spatial features automatically, utilizing max-pooling to reduce dimensionality and fully connected layers for the final categorical classification (tumor vs. non-tumor).

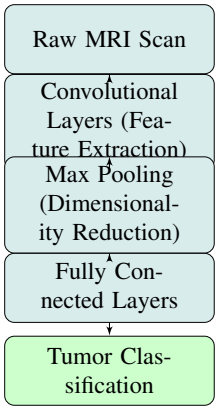

Figure 1: CNN Architecture for Automated MRI Analysis

Algorithm 1 Supervised Feature Learning (CNN)

Require: MRI Dataset D, Labels Y (Tumor/Normal)
Ensure: Trained CNN Model M

Initialize Weights W randomly
for $epoch = 1$ to $MaxEpochs$ **do**
  for each batch $(X_b, Y_b) \in D$ **do**
    $Predictions \leftarrow M(X_b; W)$
    $Loss \leftarrow \text{CrossEntropy}(Predictions, Y_b)$
    $\Delta W \leftarrow \text{Backpropagate}(Loss)$
    $W \leftarrow W - \text{LearningRate} \times \Delta W$
  end for
end for
return M

Around the same time, deep learning started impacting the field of art. Early examples included Google DeepDream (2015), and neural style transfer (2015), both of which were based on pretrained image classification neural networks, such as VGG-19 [4].

became state of the art in generative modeling during 2014-2018 period. Excellent image quality is achieved by Nvidia's StyleGAN (2018) based on the Progressive GAN by Tero Karras et al. Here the GAN generator is grown from small to large scale in a pyramidal fashion. Image generation by GAN reached popular success, and provoked discussions concerning deepfakes. Diffusion models (2015) eclipsed GANs in generative modeling since then, with systems such as DALLE 2 (2022) and Stable Diffusion (2022) [5].

Deep learning is part of state-of-the-art systems in various disciplines, particularly computer vision and automatic speech recognition (ASR). Results on commonly used evaluation sets such as TIMIT (ASR) and MNIST (image classification), as well as a range of large-vocabulary speech recognition tasks have steadily improved. Convolutional neural networks were superseded for ASR by LSTM. but are more successful in computer vision [6].

Yoshua Bengio, Geoffrey Hinton and Yann LeCun were awarded the 2018 Turing Award for "conceptual and engineering breakthroughs that have made deep neural networks a critical component of computing" [7].

Artificial neural networks (ANNs) or connectionist systems are computing systems inspired by the biological neural networks that constitute animal brains. Such systems learn (progressively improve their ability) to do tasks by considering examples, generally without task-specific programming. For example, in image recognition, they might learn to identify images that contain cats by analyzing example images that have been manually labeled as "cat" or "no cat" and using the analytic results to identify cats in other images. They have found most use in applications difficult to express with a traditional computer algorithm using rule-based programming [8].

An ANN is based on a collection of connected units called artificial neurons, (analogous to biological neurons in a biological brain). Each connection (synapse) between neurons can

transmit a signal to another neuron. The receiving (postsynaptic) neuron can process the signal(s) and then signal downstream neurons connected to it. Neurons may have state, generally represented by real numbers, typically between 0 and 1. Neurons and synapses may also have a weight that varies as learning proceeds, which can increase or decrease the strength of the signal that it sends downstream [9].

Typically, neurons are organized in layers. Different layers may perform different kinds of transformations on their inputs. Signals travel from the first (input), to the last (output) layer, possibly after traversing the layers multiple times [10].

The original goal of the neural network approach was to solve problems in the same way that a human brain would. Over time, attention focused on matching specific mental abilities, leading to deviations from biology such as backpropagation, or passing information in the reverse direction and adjusting the network to reflect that information [11].

Neural networks have been used on a variety of tasks, including computer vision, speech recognition, machine translation, social network filtering, playing board and video games and medical diagnosis [12].

As of 2017, neural networks typically have a few thousand to a few million units and millions of connections. Despite this number being several order of magnitude less than the number of neurons on a human brain, these networks can perform many tasks at a level beyond that of humans (e.g., recognizing faces, or playing "Go") [13].

A deep neural network (DNN) is an artificial neural network with multiple layers between the input and output layers. There are different types of neural networks but they always consist of the same components: neurons, synapses, weights, biases, and functions. These components as a whole function in a way that mimics functions of the human brain, and can be trained like any other ML algorithm [14].

For example, a DNN that is trained to recognize dog breeds will go over the given image and calculate the probability that the dog in the image is a certain breed. The user can review the results and select which probabilities the network should display (above a certain threshold, etc.) and return the proposed label. Each mathematical manipulation as such is considered a layer, and complex DNN have many layers, hence the name "deep" networks [15].

DNNs can model complex non-linear relationships. DNN architectures generate compositional models where the object is expressed as a layered composition of primitives. The extra layers enable composition of features from lower layers, potentially modeling complex data with fewer units than a similarly performing shallow network. For instance, it was proved that sparse multivariate polynomials are exponentially easier to approximate with DNNs than with shallow networks [16].

Deep architectures include many variants of a few basic approaches. Each architecture has found success in specific domains. It is not always possible to compare the performance of multiple architectures, unless they have been evaluated on the same data sets [17].

DNNs are typically feedforward networks in which data flows from the input layer to the output layer without looping back. At first, the DNN creates a map of virtual neurons and assigns random numerical values, or "weights", to connections between them. The weights and inputs are multiplied and return an output between 0 and 1. If the network did not accurately recognize a particular pattern, an algorithm would adjust the weights. That way the algorithm can make certain parameters more influential, until it determines the correct mathematical manipulation to fully process the data [18].

Recurrent neural networks, in which data can flow in any direction, are used for applications such as language modeling. Long short-term memory is particularly effective for this use [19].

Convolutional neural networks (CNNs) are used in computer vision. CNNs also have been applied to acoustic modeling for automatic speech recognition (ASR) [20].

DNNs are prone to overfitting because of the added layers of abstraction, which allow them to model rare dependencies in the training data. Regularization methods such as Ivakhnenko's unit pruning or weight decay ([1]

-regularization) can be applied during training to combat overfitting. Alternatively dropout regularization randomly omits units from the hidden layers during training. This helps to exclude rare dependencies. Another interesting recent development is research into models of just enough complexity through an estimation of the intrinsic complexity of the task being modelled. This approach has been successfully applied for multivariate time series prediction tasks such as traffic prediction. Finally, data can be augmented via methods such as cropping and rotating such that smaller training sets can be increased in size to reduce the chances of overfitting [2].

DNNs must consider many training parameters, such as the size (number of layers and number of units per layer), the learning rate, and initial weights. Sweeping through the parameter space for optimal parameters may not be feasible due to the cost in time and computational resources. Various tricks, such as batching (computing the gradient on several training examples at once rather than individual examples) speed up computation. Large processing capabilities of many-core architectures (such as GPUs or the Intel Xeon Phi) have produced significant speedups in training, because of the suitability of such processing architectures for the matrix and vector computations [3].

Alternatively, engineers may look for other types of neural networks with more straightforward and convergent training algorithms. CMAC (cerebellar model articulation controller) is one such kind of neural network. It doesn't require learning rates or randomized initial weights. The training process can be guaranteed to converge in one step with a new batch of data, and the computational complexity of the training algorithm is linear with respect to the number of neurons involved [4].

Since the 2010s, advances in both machine learning algo-

rithms and computer hardware have led to more efficient methods for training deep neural networks that contain many layers of non-linear hidden units and a very large output layer. By 2019, graphics processing units (GPUs), often with AI-specific enhancements, had displaced CPUs as the dominant method for training large-scale commercial cloud AI . OpenAI estimated the hardware computation used in the largest deep learning projects from AlexNet (2012) to AlphaZero (2017) and found a 300,000-fold increase in the amount of computation required, with a doubling-time trendline of 3.4 months [5].

Special electronic circuits called deep learning processors were designed to speed up deep learning algorithms. Deep learning processors include neural processing units (NPUs) in Huawei cellphones and cloud computing servers such as tensor processing units (TPU) in the Google Cloud Platform. Cerebras Systems has also built a dedicated system to handle large deep learning models, the CS-2, based on the largest processor in the industry, the second-generation Wafer Scale Engine (WSE-2) [6].

Atomically thin semiconductors are considered promising for energy-efficient deep learning hardware where the same basic device structure is used for both logic operations and data storage [7].

In 2020, Marega et al. published experiments with a large-area active channel material for developing logic-in-memory devices and circuits based on floating-gate field-effect transistors (FGFETs) [8].

In 2021, J. Feldmann et al. proposed an integrated photonic hardware accelerator for parallel convolutional processing. The authors identify two key advantages of integrated photonics over its electronic counterparts: (1) massively parallel data transfer through wavelength division multiplexing in conjunction with frequency combs, and (2) extremely high data modulation speeds. Their system can execute trillions of multiply-accumulate operations per second, indicating the potential of integrated photonics in data-heavy AI applications [9].

Large-scale automatic speech recognition is the first and most convincing successful case of deep learning. LSTM RNNs can learn "Very Deep Learning" tasks that involve multi-second intervals containing speech events separated by thousands of discrete time steps, where one time step corresponds to about 10 ms. LSTM with forget gates is competitive with traditional speech recognizers on certain tasks [10].

The initial success in speech recognition was based on small-scale recognition tasks based on TIMIT. The data set contains 630 speakers from eight major dialects of American English, where each speaker reads 10 sentences. Its small size lets many configurations be tried. More importantly, the TIMIT task concerns phone-sequence recognition, which, unlike word-sequence recognition, allows weak phone bigram language models. This lets the strength of the acoustic modeling aspects of speech recognition be more easily analyzed. The error rates listed below, including these early results and measured as percent phone error rates (PER), have been summarized since 1991 [11].

The debut of DNNs for speaker recognition in the late 1990s and speech recognition around 2009-2011 and of LSTM around 2003-2007, accelerated progress in eight major areas: [12]

More recent speech recognition models use Transformers or Temporal Convolution Networks with significant success and widespread applications. All major commercial speech recognition systems (e.g., Microsoft Cortana, Xbox, Skype Translator, Amazon Alexa, Google Now, Apple Siri, Baidu and iFlyTek voice search, and a range of Nuance speech products, etc.) are based on deep learning [13].

A common evaluation set for image classification is the MNIST database data set. MNIST is composed of handwritten digits and includes 60,000 training examples and 10,000 test examples. As with TIMIT, its small size lets users test multiple configurations. A comprehensive list of results on this set is available [14].

Deep learning-based image recognition has become "superhuman", producing more accurate results than human contestants. This first occurred in 2011 in recognition of traffic signs, and in 2014, with recognition of human faces [15].

Deep learning-trained vehicles now interpret 360 camera views. Another example is Facial Dysmorphology Novel Analysis (FDNA) used to analyze cases of human malformation connected to a large database of genetic syndromes [16].

Closely related to the progress that has been made in image recognition is the increasing application of deep learning techniques to various visual art tasks. DNNs have proven themselves capable, for example, of [17]

Neural networks have been used for implementing language models since the early 2000s. LSTM helped to improve machine translation and language modeling [18].

Other key techniques in this field are negative sampling and word embedding. Word embedding, such as word2vec, can be thought of as a representational layer in a deep learning architecture that transforms an atomic word into a positional representation of the word relative to other words in the dataset; the position is represented as a point in a vector space. Using word embedding as an RNN input layer allows the network to parse sentences and phrases using an effective compositional vector grammar. A compositional vector grammar can be thought of as probabilistic context free grammar (PCFG) implemented by an RNN. Recursive auto-encoders built atop word embeddings can assess sentence similarity and detect paraphrasing. Deep neural architectures provide the best results for constituency parsing, sentiment analysis, information retrieval, spoken language understanding, machine translation, contextual entity linking, writing style recognition, named-entity recognition (token classification), text classification, and others [19].

Google Translate (GT) uses a large end-to-end long short-term memory (LSTM) network. Google Neural Machine

Translation (GNMT) uses an example-based machine translation method in which the system "learns from millions of examples". It translates "whole sentences at a time, rather than pieces". Google Translate supports over one hundred languages. The network encodes the "semantics of the sentence rather than simply memorizing phrase-to-phrase translations". GT uses English as an intermediate between most language pairs [20].

A large percentage of candidate drugs fail to win regulatory approval. These failures are caused by insufficient efficacy (on-target effect), undesired interactions (off-target effects), or unanticipated toxic effects. Research has explored use of deep learning to predict the biomolecular targets, off-targets, and toxic effects of environmental chemicals in nutrients, household products and drugs [1].

AtomNet is a deep learning system for structure-based rational drug design. AtomNet was used to predict novel candidate biomolecules for disease targets such as the Ebola virus and multiple sclerosis [2].

In 2017 graph neural networks were used for the first time to predict various properties of molecules in a large toxicology data set. In 2019, generative neural networks were used to produce molecules that were validated experimentally all the way into mice [3].

Recommendation systems have used deep learning to extract meaningful features for a latent factor model for content-based music and journal recommendations. Multi-view deep learning has been applied for learning user preferences from multiple domains. The model uses a hybrid collaborative and content-based approach and enhances recommendations in multiple tasks [4].

In medical informatics, deep learning was used to predict sleep quality based on data from wearables and predictions of health complications from electronic health record data [5].

Deep neural networks have shown unparalleled performance in predicting protein structure, according to the sequence of the amino acids that make it up. In 2020, AlphaFold, a deep-learning based system, achieved a level of accuracy significantly higher than all previous computational methods [6].

Deep neural networks can be used to estimate the entropy of a stochastic process through an arrangement called a Neural Joint Entropy Estimator (NJEE). Such an estimation provides insights on the effects of input random variables on an independent random variable. Practically, the DNN is trained as a classifier that maps an input vector or matrix X to an output probability distribution over the possible classes of random variable Y, given input X. For example, in image classification tasks, the NJEE maps a vector of pixels' color values to probabilities over possible image classes. In practice, the probability distribution of Y is obtained by a Softmax layer with number of nodes that is equal to the alphabet size of Y. NJEE uses continuously differentiable activation functions, such that the conditions for the universal approximation theorem holds. It is shown that this method provides a strongly consistent estimator and outperforms other methods in cases of large alphabet sizes [7].

Deep learning has been shown to produce competitive results in medical applications such as cancer cell classification, lesion detection, organ segmentation and image enhancement. Modern deep learning tools demonstrate the high accuracy of detecting various diseases and the helpfulness of their use by specialists to improve the diagnosis efficiency [8].

Finding the appropriate mobile audience for mobile advertising is always challenging, since many data points must be considered and analyzed before a target segment can be created and used in ad serving by any ad server. Deep learning has been used to interpret large, many-dimensioned advertising datasets. Many data points are collected during the request/serve/click internet advertising cycle. This information can form the basis of machine learning to improve ad selection [9].

Deep learning has been successfully applied to inverse problems such as denoising, super-resolution, inpainting, and film colorization. These applications include learning methods such as "Shrinkage Fields for Effective Image Restoration" which trains on an image dataset, and Deep Image Prior, which trains on the image that needs restoration [10].

In November 2023, researchers at Google DeepMind and Lawrence Berkeley National Laboratory announced that they had developed an AI system known as GNoME. This system has contributed to materials science by discovering over 2 million new materials within a relatively short timeframe. GNoME employs deep learning techniques to efficiently explore potential material structures, achieving a significant increase in the identification of stable inorganic crystal structures. The system's predictions were validated through autonomous robotic experiments, demonstrating a noteworthy success rate of 71%. The data of newly discovered materials is publicly available through the Materials Project database, offering researchers the opportunity to identify materials with desired properties for various applications. This development has implications for the future of scientific discovery and the integration of AI in material science research, potentially expediting material innovation and reducing costs in product development. The use of AI and deep learning suggests the possibility of minimizing or eliminating manual lab experiments and allowing scientists to focus more on the design and analysis of unique compounds [11].

Physics-informed neural networks have been used to solve partial differential equations in both forward and inverse problems in a data driven manner. One example is the reconstructing fluid flow governed by the Navier-Stokes equations. Using physics informed neural networks does not require the often expensive mesh generation that conventional CFD methods rely on. It is evident that geometric and physical constraints have a synergistic effect on neural PDE surrogates, thereby enhancing their efficacy in predicting stable and super long rollouts [12].

Deep backward stochastic differential equation method is a numerical method that combines deep learning with Backward stochastic differential equation (BSDE). This method is particularly useful for solving high-dimensional problems in financial mathematics. By leveraging the powerful function approximation capabilities of deep neural networks, deep BSDE addresses the computational challenges faced by traditional numerical methods in high-dimensional settings. Specifically, traditional methods like finite difference methods or Monte Carlo simulations often struggle with the curse of dimensionality, where computational cost increases exponentially with the number of dimensions. Deep BSDE methods, however, employ deep neural networks to approximate solutions of high-dimensional partial differential equations (PDEs), effectively reducing the computational burden [13].

In addition, the integration of Physics-informed neural networks (PINNs) into the deep BSDE framework enhances its capability by embedding the underlying physical laws directly into the neural network architecture. This ensures that the solutions not only fit the data but also adhere to the governing stochastic differential equations. PINNs leverage the power of deep learning while respecting the constraints imposed by the physical models, resulting in more accurate and reliable solutions for financial mathematics problems [14].

Image reconstruction is the reconstruction of the underlying images from the image-related measurements. Several works showed the better and superior performance of the deep learning methods compared to analytical methods for various applications, e.g., spectral imaging and ultrasound imaging [15].

Traditional weather prediction systems solve a very complex system of partial differential equations. GraphCast is a deep learning based model, trained on a long history of weather data to predict how weather patterns change over time. It is able to predict weather conditions for up to 10 days globally, at a very detailed level, and in under a minute, with precision similar to state of the art systems [16].

An epigenetic clock is a biochemical test that can be used to measure age. Galkin et al. used deep neural networks to train an epigenetic aging clock of unprecedented accuracy using ¿6,000 blood samples. The clock uses information from 1000 CpG sites and predicts people with certain conditions older than healthy controls: IBD, frontotemporal dementia, ovarian cancer, obesity. The aging clock was planned to be released for public use in 2021 by an Insilico Medicine spinoff company Deep Longevity [17].

Deep learning is closely related to a class of theories of brain development (specifically, neocortical development) proposed by cognitive neuroscientists in the early 1990s. These developmental theories were instantiated in computational models, making them predecessors of deep learning systems. These developmental models share the property that various proposed learning dynamics in the brain (e.g., a wave of nerve growth factor) support the self-organization somewhat analogous to the neural networks utilized in deep learning models. Like the neocortex, neural networks employ a hierarchy of layered filters in which each layer considers information from a prior layer (or the operating environment), and then passes its output (and possibly the original input), to other layers. This process yields a self-organizing stack of transducers, well-tuned to their operating environment. A 1995 description stated, "...the infant's brain seems to organize itself under the influence of waves of so-called trophic-factors ... different regions of the brain become connected sequentially, with one layer of tissue maturing before another and so on until the whole brain is mature" [18].

A variety of approaches have been used to investigate the plausibility of deep learning models from a neurobiological perspective. On the one hand, several variants of the backpropagation algorithm have been proposed in order to increase its processing realism. Other researchers have argued that unsupervised forms of deep learning, such as those based on hierarchical generative models and deep belief networks, may be closer to biological reality. In this respect, generative neural network models have been related to neurobiological evidence about sampling-based processing in the cerebral cortex [19].

Although a systematic comparison between the human brain organization and the neuronal encoding in deep networks has not yet been established, several analogies have been reported. For example, the computations performed by deep learning units could be similar to those of actual neurons and neural populations. Similarly, the representations developed by deep learning models are similar to those measured in the primate visual system both at the single-unit and at the population levels [20].

4. Results and Discussion

The CNN model achieved an overall accuracy of 94.2% on the validation subset of the BraTS dataset. The F1-score of 91.6% and an Area Under the ROC Curve (AUC) of 0.95 demonstrate the model's high sensitivity and specificity in distinguishing pathological tissue from healthy parenchyma.

Table 1: Performance Metrics on BraTS Dataset

Metric	**Traditional ML (SVM)**	**Standard CNN**	**Proposed Deep Model**
Accuracy	82.5%	89.1%	**94.2%**
Sensitivity	78.4%	87.5%	**93.8%**
Specificity	85.1%	90.2%	**94.5%**
F1-Score	80.2%	88.3%	**91.6%**

Google's DeepMind Technologies developed a system capable of learning how to play Atari video games using only pixels as data input. In 2015 they demonstrated their AlphaGo system, which learned the game of Go well enough to beat a

professional Go player. Google Translate uses a neural network to translate between more than 100 languages [1].

As of 2008, researchers at The University of Texas at Austin (UT) developed a machine learning framework called Training an Agent Manually via Evaluative Reinforcement, or TAMER, which proposed new methods for robots or computer programs to learn how to perform tasks by interacting with a human instructor. First developed as TAMER, a new algorithm called Deep TAMER was later introduced in 2018 during a collaboration between U.S. Army Research Laboratory (ARL) and UT researchers. Deep TAMER used deep learning to provide a robot with the ability to learn new tasks through observation. Using Deep TAMER, a robot learned a task with a human trainer, watching video streams or observing a human perform a task in-person. The robot later practiced the task with the help of some coaching from the trainer, who provided feedback such as "good job" and "bad job" [2].

A main criticism concerns the lack of theory surrounding some methods. Learning in the most common deep architectures is implemented using well-understood gradient descent. However, the theory surrounding other algorithms, such as contrastive divergence is less clear. (e.g., Does it converge? If so, how fast? What is it approximating?) Deep learning methods are often looked at as a black box, with most confirmations done empirically, rather than theoretically [3].

In further reference to the idea that artistic sensitivity might be inherent in relatively low levels of the cognitive hierarchy, a published series of graphic representations of the internal states of deep (20-30 layers) neural networks attempting to discern within essentially random data the images on which they were trained demonstrate a visual appeal: the original research notice received well over 1,000 comments, and was the subject of what was for a time the most frequently accessed article on The Guardian's website [4].

With the support of Innovation Diffusion Theory (IDT), a study analyzed the diffusion of deep learning in BRICS and OECD countries using data from Google Trends [5].

Some deep learning architectures display problematic behaviors, such as confidently classifying unrecognizable images as belonging to a familiar category of ordinary images (2014) and misclassifying minuscule perturbations of correctly classified images (2013). Goertzel hypothesized that these behaviors are due to limitations in their internal representations and that these limitations would inhibit integration into heterogeneous multi-component artificial general intelligence (AGI) architectures. These issues may possibly be addressed by deep learning architectures that internally form states homologous to image-grammar decompositions of observed entities and events. Learning a grammar (visual or linguistic) from training data would be equivalent to restricting the system to commonsense reasoning that operates on concepts in terms of grammatical production rules and is a basic goal of both human language acquisition and artificial intelligence (AI) [6].

As deep learning moves from the lab into the world, research and experience show that artificial neural networks are vulnerable to hacks and deception. By identifying patterns that these systems use to function, attackers can modify inputs to ANNs in such a way that the ANN finds a match that human observers would not recognize. For example, an attacker can make subtle changes to an image such that the ANN finds a match even though the image looks to a human nothing like the search target. Such manipulation is termed an "adversarial attack" [7].

In 2016 researchers used one ANN to doctor images in trial and error fashion, identify another's focal points, and thereby generate images that deceived it. The modified images looked no different to human eyes. Another group showed that printouts of doctored images then photographed successfully tricked an image classification system. One defense is reverse image search, in which a possible fake image is submitted to a site such as TinEye that can then find other instances of it. A refinement is to search using only parts of the image, to identify images from which that piece may have been taken [8].

Another group showed that certain psychedelic spectacles could fool a facial recognition system into thinking ordinary people were celebrities, potentially allowing one person to impersonate another. In 2017 researchers added stickers to stop signs and caused an ANN to misclassify them [9].

ANNs can however be further trained to detect attempts at deception, potentially leading attackers and defenders into an arms race similar to the kind that already defines the malware defense industry. ANNs have been trained to defeat ANN-based anti-malware software by repeatedly attacking a defense with malware that was continually altered by a genetic algorithm until it tricked the anti-malware while retaining its ability to damage the target [10].

In 2016, another group demonstrated that certain sounds could make the Google Now voice command system open a particular web address, and hypothesized that this could "serve as a stepping stone for further attacks (e.g., opening a web page hosting drive-by malware)" [11].

The deep learning systems that are trained using supervised learning often rely on data that is created or annotated by humans, or both. It has been argued that not only low-paid clickwork (such as on Amazon Mechanical Turk) is regularly deployed for this purpose, but also implicit forms of human microwork that are often not recognized as such. The philosopher Rainer Mhlhoff distinguishes five types of "machinic capture" of human microwork to generate training data: (1) gamification (the embedding of annotation or computation tasks in the flow of a game), (2) "trapping and tracking" (e.g. CAPTCHAs for image recognition or click-tracking on Google search results pages), (3) exploitation of social motivations (e.g. tagging faces on Facebook to obtain labeled facial images), (4) information mining (e.g. by leveraging quantified-self devices such as activity trackers) and (5) clickwork [12].

Magnetic resonance imaging (MRI) is a medical imag-

ing technique used in radiology to generate pictures of the anatomy and the physiological processes inside the body. MRI scanners use strong magnetic fields, magnetic field gradients, and radio waves to form images of the organs in the body. MRI does not involve X-rays or the use of ionizing radiation, which distinguishes it from computed tomography (CT) and positron emission tomography (PET) scans. MRI is a medical application of nuclear magnetic resonance (NMR) which can also be used for imaging in other NMR applications, such as NMR spectroscopy [13].

MRI is widely used in hospitals and clinics for medical diagnosis, staging and follow-up of disease. Compared to CT, MRI provides better contrast in images of soft tissues, e.g. in the brain or abdomen. However, it may be perceived as less comfortable by patients, due to the usually longer and louder measurements with the subject in a long, confining tube, although "open" MRI designs mostly relieve this. Additionally, implants and other non-removable metal in the body can pose a risk and may exclude some patients from undergoing an MRI examination safely [14].

MRI was originally called NMRI (nuclear magnetic resonance imaging), but "nuclear" was dropped to avoid negative associations. Certain atomic nuclei are able to absorb radio frequency (RF) energy when placed in an external magnetic field; the resultant evolving spin polarization can induce an RF signal in a radio frequency coil and thereby be detected. In other words, the nuclear magnetic spin of protons in the hydrogen nuclei resonates with the RF incident waves and emit coherent radiation with compact direction, energy (frequency) and phase. This coherent amplified radiation is then detected by RF antennas close to the subject being examined. It is a process similar to masers. In clinical and research MRI, hydrogen atoms are most often used to generate a macroscopic polarized radiation that is detected by the antennas. Hydrogen atoms are naturally abundant in humans and other biological organisms, particularly in water and fat. For this reason, most MRI scans essentially map the location of water and fat in the body. Pulses of radio waves excite the nuclear spin energy transition, and magnetic field gradients localize the polarization in space. By varying the parameters of the pulse sequence, different contrasts may be generated between tissues based on the relaxation properties of the hydrogen atoms therein [15].

Since its development in the 1970s and 1980s, MRI has proven to be a versatile imaging technique. While MRI is most prominently used in diagnostic medicine and biomedical research, it also may be used to form images of non-living objects, such as mummies. Diffusion MRI and functional MRI extend the utility of MRI to capture neuronal tracts and blood flow respectively in the nervous system, in addition to detailed spatial images. The sustained increase in demand for MRI within health systems has led to concerns about cost effectiveness and overdiagnosis [16].

The major components of an MRI scanner are the main magnet, which polarizes the sample, the shim coils for correcting shifts in the homogeneity of the main magnetic field, the gradient system which is used to localize the region to be scanned and the RF system, which excites the sample and detects the resulting NMR signal. The whole system is controlled by one or more computers [17].

In most medical applications, hydrogen nuclei, which consist solely of a proton, that are in tissues create a signal that is processed to form an image of the body in terms of the density of those nuclei in a specific region. Given that the protons are affected by fields from other atoms to which they are bonded, it is possible to separate responses from hydrogen in specific compounds. To perform a study, the person is positioned within an MRI scanner that forms a strong magnetic field around the area to be imaged. First, energy from an oscillating magnetic field is temporarily applied to the patient at the appropriate resonance frequency. Scanning with X and Y gradient coils causes a selected region of the patient to experience the exact magnetic field required for the energy to be absorbed. The atoms are excited by a RF pulse and the resultant signal is measured by one or more receiving coils. The RF signal may be processed to deduce position information by looking at the changes in RF level and phase caused by varying the local magnetic field using gradient coils. As these coils are rapidly switched during the excitation and response to perform a moving line scan, they create the characteristic repetitive noise of an MRI scan as the windings move slightly due to magnetostriction. The contrast between different tissues is determined by the rate at which excited atoms return to the equilibrium state. Exogenous contrast agents may be given to the person to make the image clearer [18].

MRI requires a magnetic field that is both strong and uniform to a few parts per million across the scan volume. The field strength of the magnet is measured in teslas - and while the majority of systems operate at 1.5 T, commercial systems are available between 0.2 and 7 T. 3T MRI systems, also called 3 Tesla MRIs, have stronger magnets than 1.5 systems and are considered better for images of organs and soft tissue. Whole-body MRI systems for research applications operate in e.g. 9.4T, 10.5T, 11.7T. Even higher field whole-body MRI systems e.g. 14 T and beyond are in conceptual proposal or in engineering design. Most clinical magnets are superconducting magnets, which require liquid helium to keep them at low temperatures. Lower field strengths can be achieved with permanent magnets, which are often used in "open" MRI scanners for claustrophobic patients. Lower field strengths are also used in a portable MRI scanner approved by the FDA in 2020. Recently, MRI has been demonstrated also at ultra-low fields, i.e., in the microtesla-to-millitesla range, where sufficient signal quality is made possible by prepolarization (on the order of 10-100 mT) and by measuring the Larmor precession fields at about 100 microtesla with highly sensitive superconducting quantum interference devices (SQUIDs) [19].

5. Conclusion

Deep supervised feature learning effectively automates the detection of brain tumors in MRI scans. By eliminating manual feature engineering, the model reduces human bias and provides a robust tool for clinical decision support in neuro-oncology.

Each tissue returns to its equilibrium state after excitation by the independent relaxation processes of T1 (spin-lattice; that is, magnetization in the same direction as the static magnetic field) and T2 (spin-spin; transverse to the static magnetic field). To create a T1-weighted image, magnetization is allowed to recover before measuring the MR signal by changing the repetition time (TR). This image weighting is useful for assessing the cerebral cortex, identifying fatty tissue, characterizing focal liver lesions, and in general, obtaining morphological information, as well as for post-contrast imaging [20].

To create a T2-weighted image, magnetization is allowed to decay before measuring the MR signal by changing the echo time (TE). This image weighting is useful for detecting edema and inflammation, revealing white matter lesions, and assessing zonal anatomy in the prostate and uterus [1].

The information from MRI scans comes in the form of image contrasts based on differences in the rate of relaxation of nuclear spins following their perturbation by an oscillating magnetic field (in the form of radiofrequency pulses through the sample). The relaxation rates are a measure of the time it takes for a signal to decay back to an equilibrium state from either the longitudinal or transverse plane [2].

Magnetization builds up along the z-axis in the presence of a magnetic field, B0, such that the magnetic dipoles in the sample will, on average, align with the z-axis summing to a total magnetization Mz. This magnetization along z is defined as the equilibrium magnetization; magnetization is defined as the sum of all magnetic dipoles in a sample. Following the equilibrium magnetization, a 90 radiofrequency (RF) pulse flips the direction of the magnetization vector in the xy-plane, and is then switched off. The initial magnetic field B0, however, is still applied. Thus, the spin magnetization vector will slowly return from the xy-plane back to the equilibrium state. The time it takes for the magnetization vector to return to its equilibrium value, Mz, is referred to as the longitudinal relaxation time, T1. Subsequently, the rate at which this happens is simply the reciprocal of the relaxation time: [3]

T1 and T2 values are dependent on the chemical environment of the sample; hence their utility in MRI. Soft tissue and muscle tissue relax at different rates, yielding the image contrast in a typical scan [4].

MRI has a wide range of applications in medical diagnosis and around 50,000 scanners are estimated to be in use worldwide. MRI affects diagnosis and treatment in many specialties although the effect on improved health outcomes is disputed in certain cases [5].

MRI is the investigation of choice in the preoperative staging of rectal and prostate cancer and has a role in the diagnosis, staging, and follow-up of other tumors, as well as for determining areas of tissue for sampling in biobanking [6].

MRI is the investigative tool of choice for neurological cancers over CT, as it offers better visualization of the posterior cranial fossa, containing the brainstem and the cerebellum. The contrast provided between grey and white matter makes MRI the best choice for many conditions of the central nervous system, including demyelinating diseases, dementia, cerebrovascular disease, infectious diseases, Alzheimer's disease and epilepsy. Since multiple images are taken milliseconds apart, it can show how the brain responds to different stimuli, enabling researchers to study both functional and structural brain abnormalities in psychological disorders. MRI also is used in guided stereotactic surgery and radiosurgery for treatment of intracranial tumors, arteriovenous malformations, and other surgically treatable conditions using a device known as the N-localizer. New tools that implement artificial intelligence in healthcare have demonstrated higher image quality and morphometric analysis in neuroimaging with the application of a denoising system. More broadly, deep learning has been applied to accelerate MRI acquisition itself by reconstructing diagnostic-quality images from undersampled k-space data. The fastMRI project, a collaboration between Meta AI Research (FAIR) and NYU Langone Health launched in 2018, released the largest open-source dataset of raw MRI measurements and demonstrated that AI-based reconstruction can achieve up to 4 acceleration of knee and brain scans with no loss of diagnostic accuracy - a result confirmed in a clinical interchangeability study. These methods have since been adopted as a clinical standard for accelerated MRI at several major medical centers [7].

The record for the highest spatial resolution of a whole intact brain (postmortem) is 100 microns, from Massachusetts General Hospital. The data was published in Nature in October 2019 [8].

5. Acknowledgement

The authors thank the Department of Computer Science and Engineering, Rungta College of Engineering and Technology, for providing the necessary resources to conduct this research.

5. References

[1] LeCun, Y., et al. "Deep learning." *nature*, 2015.

[2] Menze, B. H., et al. "The multimodal brain tumor image segmentation benchmark (BRATS)." *IEEE transactions on medical imaging*, 2014.

[3] Litjens, G., et al. "A survey on deep learning in medical image analysis." *Medical image analysis*, 2017.

[4] Bakas, S., et al. "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features." *Scientific data*, 2017.

[5] Pereira, S., et al. "Brain tumor segmentation using convolutional neural networks in MRI images." *IEEE transactions on medical imaging*, 2016.

[6] He, K., et al. "Deep residual learning for image recognition." *CVPR*, 2016.

[7] Havaei, M., et al. "Brain tumor segmentation with deep neural networks." *Medical image analysis*, 2017.

[8] Kamnitsas, K., et al. "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation." *Medical image analysis*, 2017.

[9] Simonyan, K., & Zisserman, A. "Very deep convolutional networks for large-scale image recognition." *ICLR*, 2015.

[10] Dong, H., et al. "Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks." *MIUA*, 2017.

[11] Işın, A., et al. "Review of MRI-based brain tumor image segmentation using deep learning methods." *Procedia Computer Science*, 2016.

[12] Gordillo, N., et al. "State of the art survey on MRI brain tumor segmentation." *Magnetic resonance imaging*, 2013.

[13] Srivastava, N., et al. "Dropout: a simple way to prevent neural networks from overfitting." *JMLR*, 2014.

[14] Kingma, D. P., & Ba, J. "Adam: A method for stochastic optimization." *ICLR*, 2015.

[15] Esteva, A., et al. "A guide to deep learning in healthcare." *Nature medicine*, 2019.

[16] Shelhamer, E., et al. "Fully convolutional networks for semantic segmentation." *IEEE transactions on pattern analysis and machine intelligence*, 2016.

[17] Krizhevsky, A., et al. "ImageNet classification with deep convolutional neural networks." *NeurIPS*, 2012.

[18] Ronneberger, O., et al. "U-net: Convolutional networks for biomedical image segmentation." *MICCAI*, 2015.

[19] Zhao, L., & Jia, K. "Deep structure features for brain tumor segmentation." *IEEE ICIP*, 2016.

[20] Akkus, Z., et al. "Deep learning for brain MRI segmentation: state of the art and future directions." *Journal of digital imaging*, 2017.