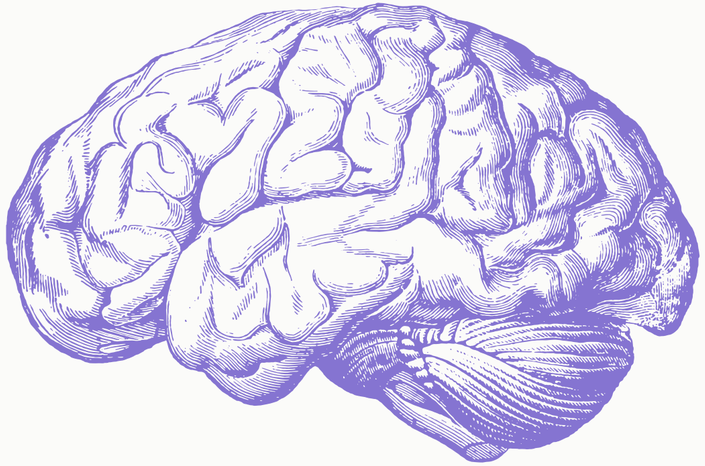
9TH INTERNATIONAL CONFERENCE ON COMPUTER SCIENCE AND COMPUTATIONAL INTELLIGENCE 2024 (ICCSCI 2024)

IDENTIFICATION OF BIOMARKERS FOR BRAIN CANCER:

INTEGRATING BIOINFORMATICS AND DEEP LEARNING

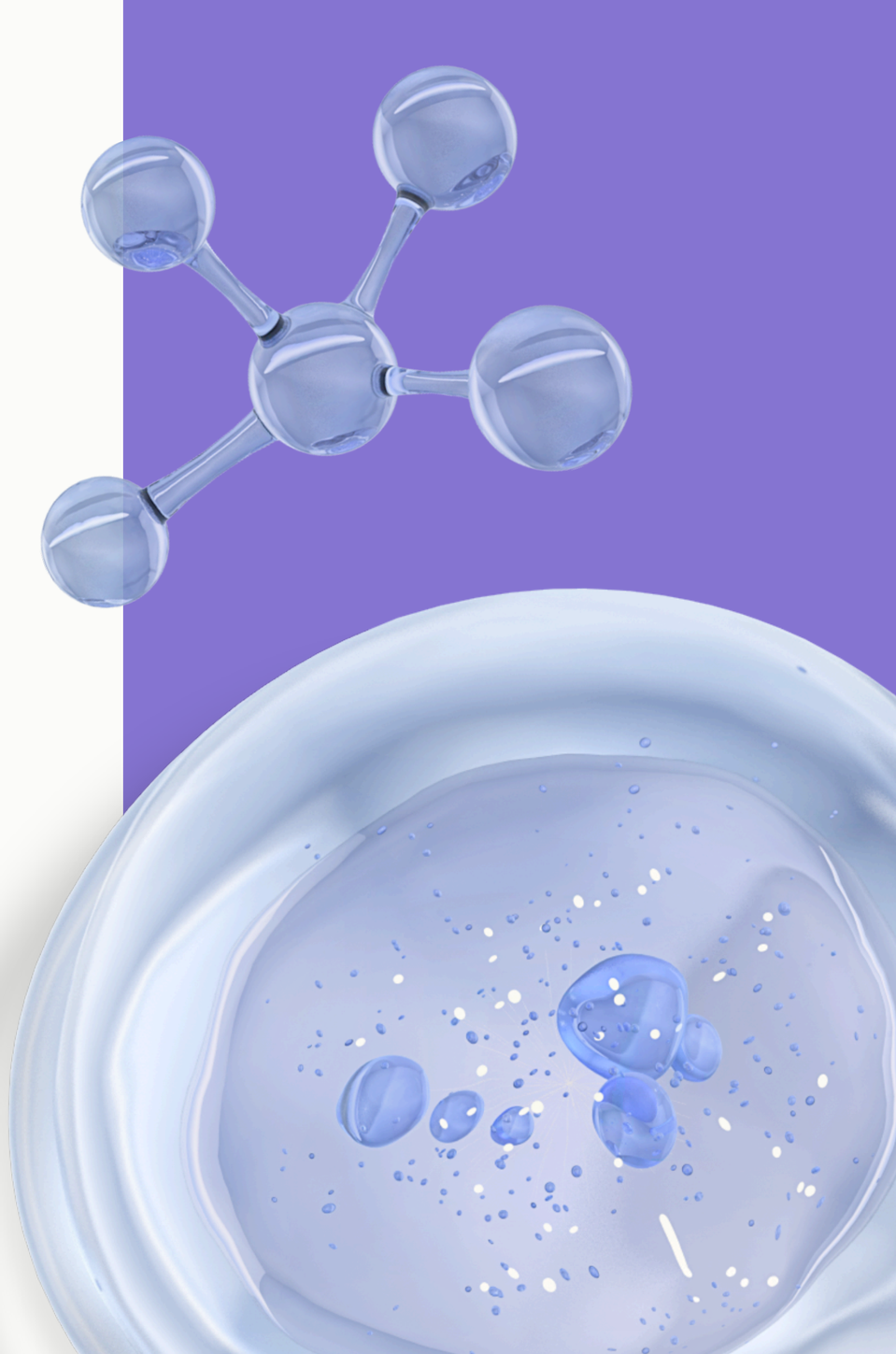


AICC - Bioinformatics

Jonathan Alvindo Fernandi
Vinson Luckianto
Ghinaa Zain Nabiilah
Jurike V. Moniaga

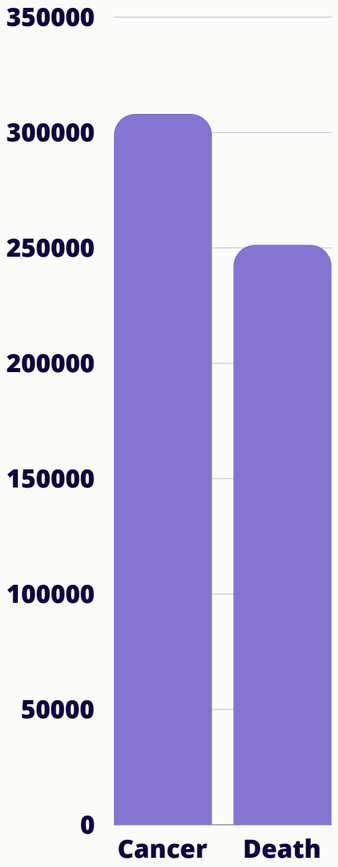
TABLE
OF CONTENTS

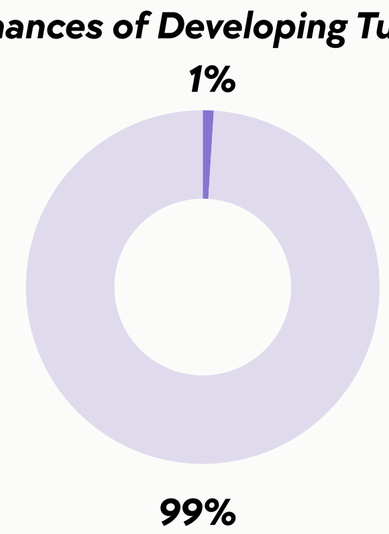
Introduction



Introduction

1. Prevalence of Brain Cancer
2. Importance of Early Detection
3. Roles of MRI in Brain Tumor Diagnosis
4. Challenges in Detecting Brain Cancer using MRI





Brain Tumors Data

- 308,102 people were diagnosed with brain tumors in 2020 and 251,329 people died from brain cancer & CNS tumors
- A person's lifetime chance of developing a tumor is less than 1%
- The 10th leading cause of death for men & women

Data gathered from the research website: Cancer.net

5W + 1H

- **What?**

  What is the research that we want to address?

- **Where?**

  Where can the supporting research be found?

- **Why?**

  Why did we come up with this research topic?

- **When?**

  When did research on brain cancer and tumors begin?

- **Who?**

  Who will this research benefit?

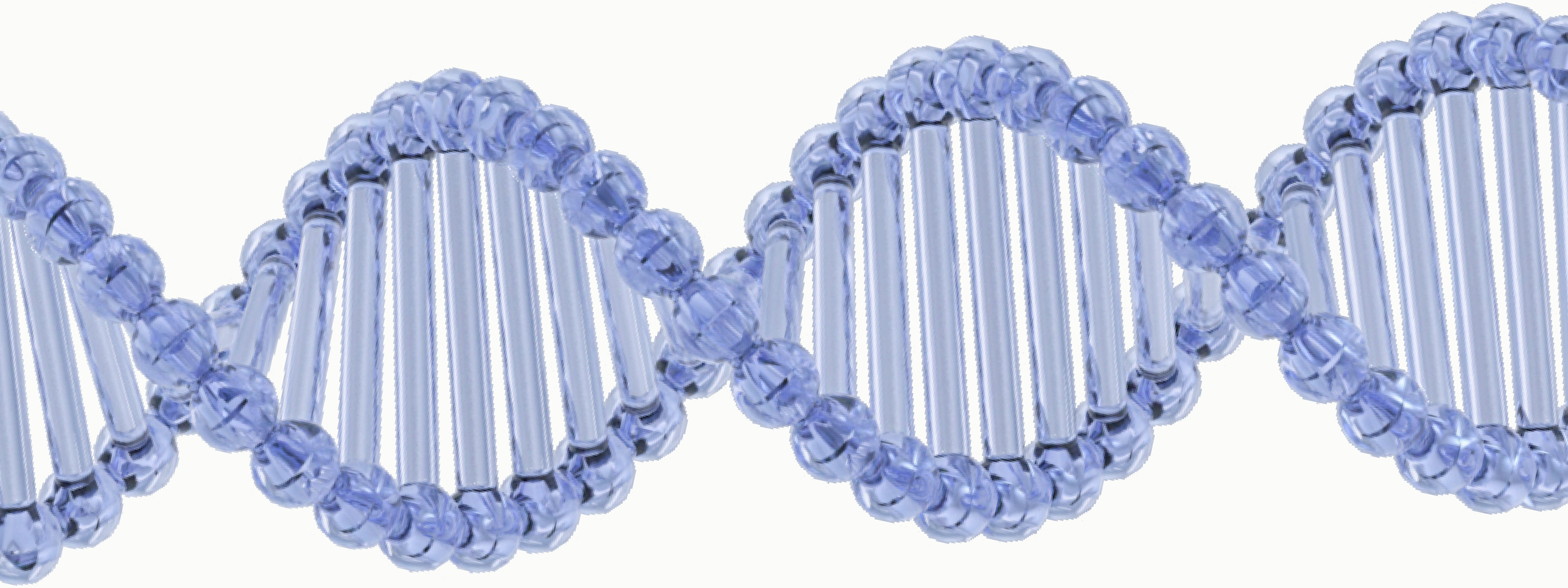
- **How?**

  How is the research methodology implemented?

DATASET

The dataset used in this research was gathered from the Kaggle website: "Brain Tumor MRI Dataset"

https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset



Related Works

Related Works

Related Paper

Pei et al. (2020) CANet: Unified Brain Cancer Analysis Framework

Extended CANet CNN architecture

Jointly performs tumor segmentation, classification, and survival prediction

Demonstrates potential of integrating multiple tasks for diagnosis and prognosis

Qureshi et al. (2023) Radiogenomic Model for Glioblastoma Biomarker Prediction

Fuses multi-parametric MRI (mpMRI) and genomic data

Predicts MGMT promoter methylation status, a key biomarker

Achieves AUROC of 0.92, showcasing the power of combining imaging and molecular data with deep learning

Tabassum et al. (2023) Review of Radiomics and Machine Learning for Brain Tumor Classification

Highlights importance of CNNs and integrating radiomic features with multi-omics data

Emphasizes potential of multi-omics data integration with deep learning

Aims to identify novel biomarkers for brain cancer

Methods

Methods

Data Preprocessing

- Dataset: 7023 brain MRI scans from Kaggle
- Split into training and testing subsets using stratified sampling
- Testing subset further divided into validation and test groups
- Pandas DataFrames created with file paths and class labels
- Data visualization using Seaborn countplots to analyze class distributions
- Keras ImageDataGenerator used for preprocessing
- Data augmentation (rescaling, brightness adjustment) applied to training and validation sets
- Testing set only rescaled

Model Building

- Pre-trained Xception on ImageNet used as base model for transfer learning
- Top layers removed, remaining layers frozen
- Additional layers added:
  - Flatten layer
  - Dropout layers
  - Dense layers with ReLU and Softmax activations
- Model compiled with:
  - Adamax optimizer
  - Categorical cross-entropy loss
- Evaluation metrics: accuracy, precision, recall

Methods

Training

- Model trained for 10 epochs using Keras "fit" function
- Training and validation data generators used as inputs
- Model weights adjusted based on validation set performance
- Matplotlib used to visualize training and validation metrics over epochs
- Metrics: loss, accuracy, precision, recall
- Best epoch identified for each metric

Evaluation

- Loss and evaluation metrics calculated for training, validation, and testing datasets using Keras "evaluate" function
- Model predictions on testing set printed using "predict" function
- Confusion matrix visualized using Seaborn
- Classification report generated with performance metrics for each class

Testing & Development

- Trained model saved in HDF5 format using Keras "save" function
- Model easily loaded and deployed using "load_model" function
- Custom "description" function provides detailed information about predicted tumor type
- Overview, symptoms, treatments, relevant specialists
- "predict" function loads and preprocesses MRI image, generates predictions, displays probability distribution, and calls "description" function
- Individual predictions showcased by invoking "predict" function on MRI images from testing set

Data Analysis & Results

Data Analysis

Model Training & Evaluation

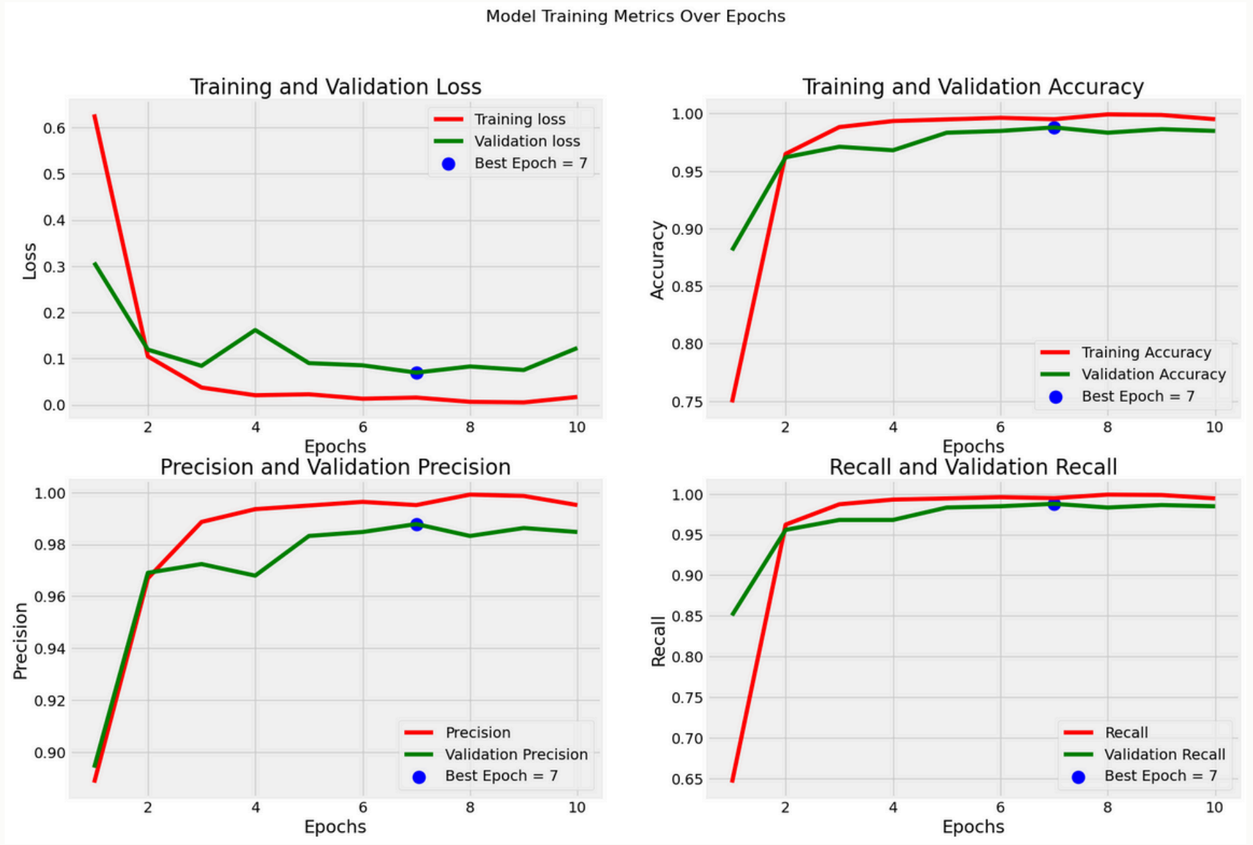


Table 1. The Loss and Evaluation Metrics

	Loss	Accuracy
Train	0.0201	99.53%
Validation	0.1201	98.47%
Test	0.0970	97.87%

Data Analysis

Model Prediction & Evaluation

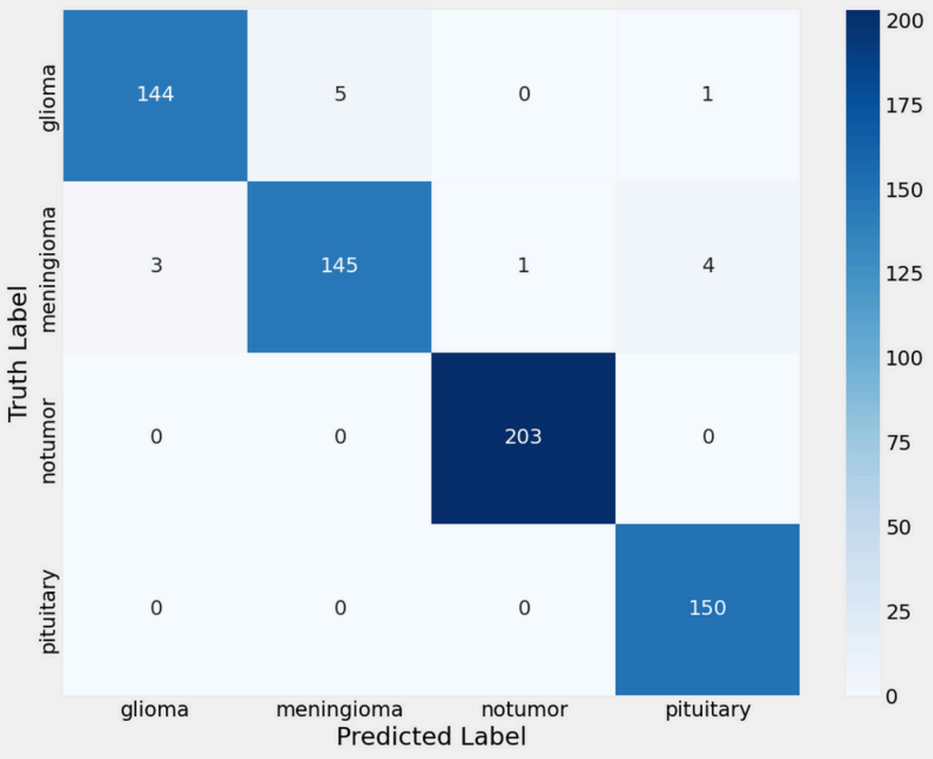
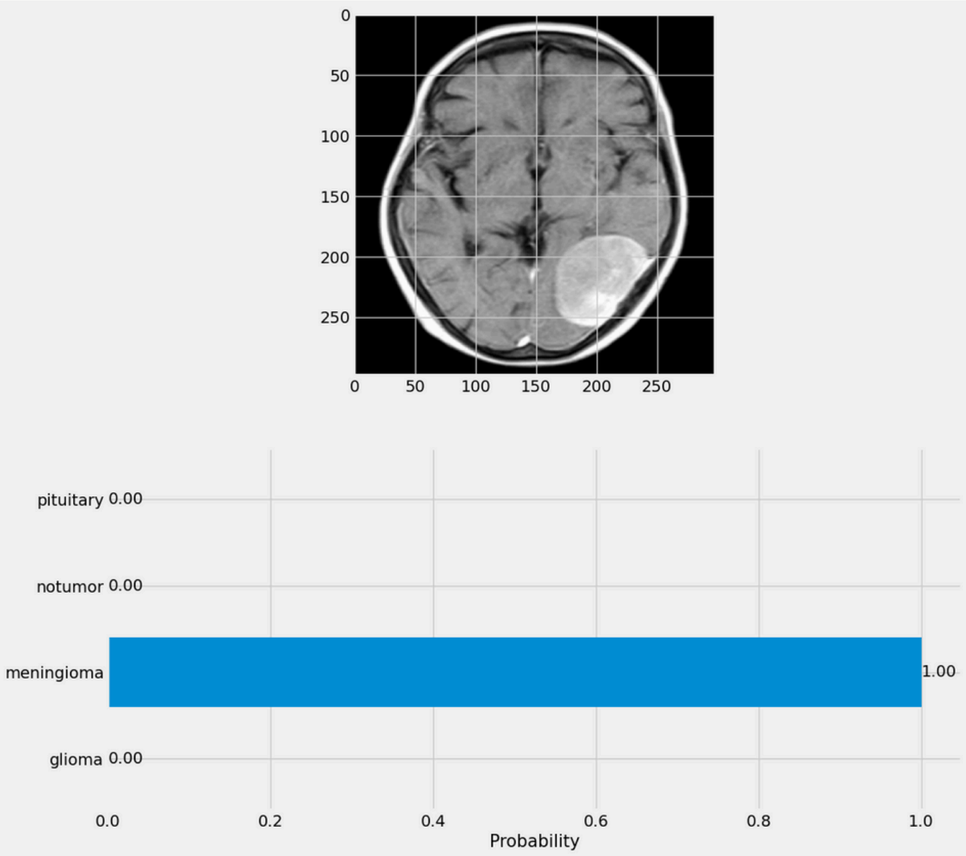


Table 2. Classification Report

	Precision	Recall	F_1 Score	Support
Glioma	0.98	0.96	0.97	150
Meningioma	0.97	0.95	0.96	153
No Tumor	1.00	1.00	1.00	203
Pituitary	0.97	1.00	0.98	150
Accuracy			0.98	656
Macro Average	0.98	0.98	0.98	656
Weighted Average	0.98	0.98	0.98	656

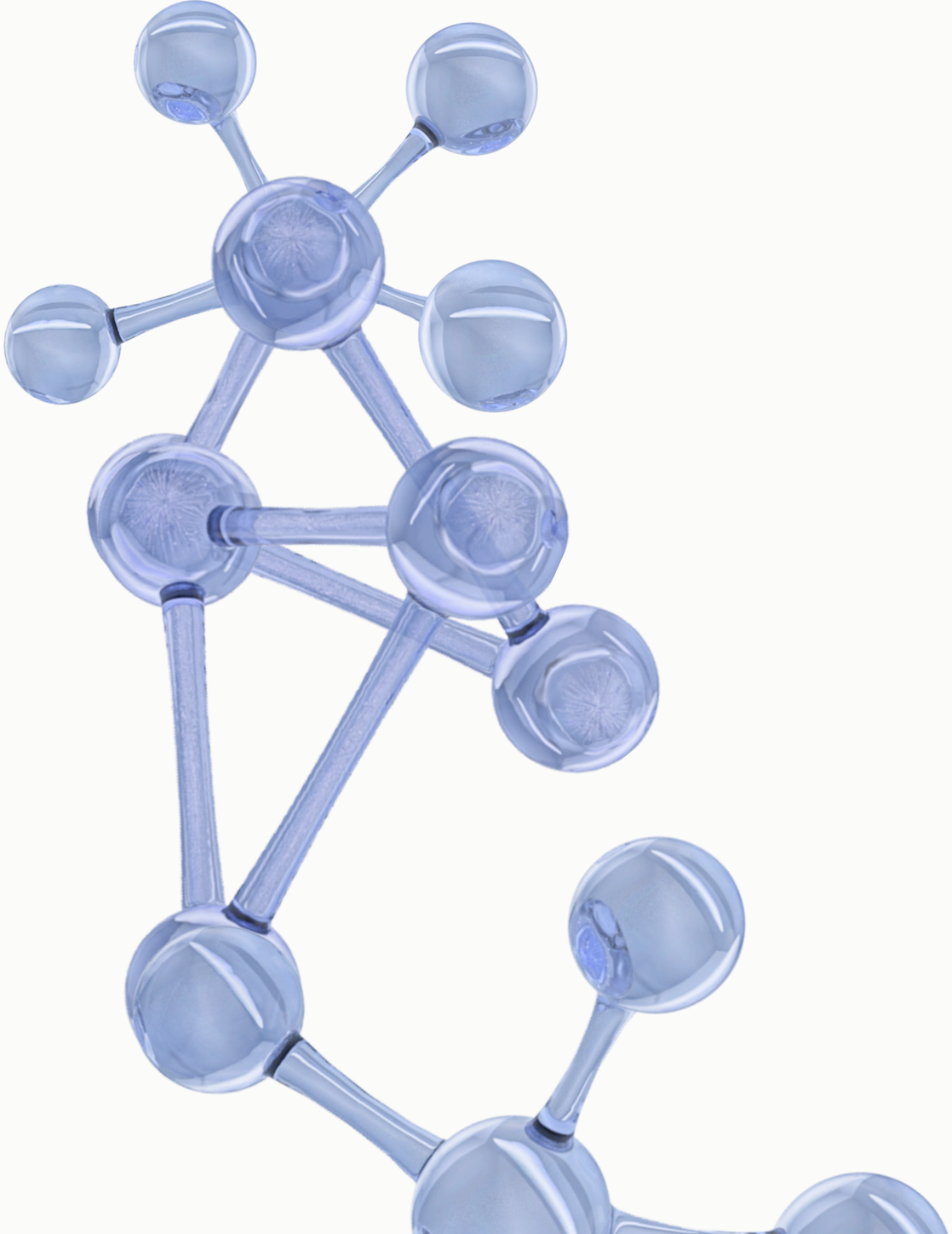
RESULTS

Example of the Prediction Bar Chart



Example of the Prediction Description

```
Predicted Tumor Type:
1. Meningioma
Probability: 100.00%

a. Overview
A usually non-cancerous tumor that arises from the membranes surrounding the brain and spinal cord.

- It is not clear what causes a meningioma. Radiation therapy, female hormones, and genetics may play a role. In most cases, the condition is non-cancerous.
- Symptoms may include changes in vision, headaches, hearing loss, and seizures.
- A small, slow-growing meningioma that is not causing signs or symptoms may not require treatment. When required, treatment might involve surgery or radiation.

[!] Rare
Fewer than 150 thousand cases per year

- Treatable by a medical professional
- Requires a medical diagnosis
- Lab tests or imaging are always required
- Chronic: can last for years or be lifelong

b. Symptoms
[!] Requires a medical diagnosis
Symptoms may include changes in vision, headaches, hearing loss, and seizures.

[!] People may experience:
- Visual: double vision or vision loss
- Eyes: swelling of the optic disc or unequal pupils
- Also common: headache, impaired coordination, or seizure


c. Treatments
Note: Treatment depends on stage
Treatments include surgery, radiation therapy, chemotherapy, and targeted molecular therapy.

[!] Supportive Care
- Monitoring
Regular follow-up to evaluate for improvement.

[!] Surgery
- Craniotomy
Brain surgery in which a piece of bone is removed from the skull.
- Microsurgery
Surgery is performed on very small blood vessels and nerves using a microscope.

[!] Medical Procedure
- Radiation Therapy
Treatment that uses X-rays and other high-energy rays to kill abnormal cells.

d. Specialists
- Neurosurgery
Specialises in nervous system disorders.
- Radiation Oncology
Treats and manages cancer by prescribing radiation therapy.
- Neurology
Treats nervous system disorders.
- Oncology
Specialises in cancer.
```



Conclusion

& Future Works

Conclusions

- Proposed a deep learning model based on Xception CNN for identifying and classifying brain tumors from MRI scans
- Achieved 97.87% accuracy on the testing set, with high precision and recall scores across all tumor types
- Model's exceptional performance in detecting the absence of tumors and pituitary tumors with perfect precision
- Demonstrates the potential for non-invasive and accurate brain tumor diagnosis using deep learning
- Combines bioinformatics techniques and deep learning algorithms for improved treatment planning and personalized medicine in neuro-oncology

Future Works

- Integrate additional imaging modalities, such as CT scans or PET images, to enhance the model's performance and provide more comprehensive information for diagnosis
- Incorporate multi-omics data, including genomic, transcriptomic, and proteomic information, to identify molecular biomarkers and gain insights into cancer biology and treatment response
- Validate the model's performance using diverse datasets of tumor types, patient demographics, and imaging protocols to ensure generalizability across different clinical settings
- Develop a unified framework that models the interplay between tumor segmentation, classification, and grading tasks
- Explore the biological basis of learned radiomic features to provide interpretability and support clinical decision-making
- Optimize the model's architecture to balance high performance with computational efficiency for real-time use in clinical settings

THANK YOU FOR YOUR ATTENTION

Jonathan Alvindo Fernandi
Vinson Luckianto
Ghinaa Zain Nabiilah
Jurike V. Moniaga