

Towards Multimodal Deep Learning for Predicting Survival of Brain Metastatic Cancer

Inês N. Carvalho ^{*1}, Carla Guerreiro², Cristiano Esteves², Matilde Calheira¹, Stefani Domentean³, Rita Cascão³, Mariana M. Santos¹, Cláudia C. Faria^{3, 4}, and Joao C. Guimaraes¹

¹Faculdade de Medicina, Universidade de Lisboa, 1649-028 Lisboa, Portugal

²Serviço de Imagiologia, Hospital de Santa Maria, Centro Hospitalar Universitário Lisboa Norte (CHULN), 1649-035 Lisboa, Portugal

³Gulbenkian Institute for Molecular Medicine (GIMM), Avenida Prof. Egas Moniz, 1649-035 Lisboa, Portugal

⁴Serviço de Neurocirurgia, Hospital de Santa Maria, Unidade Local de Saúde de Santa Maria (ULSSM), 1649-035 Lisboa, Portugal

ines.c@edu.ulisboa.pt

Abstract

Brain metastases (BM) are associated with poor prognosis and high morbidity, yet survival prediction remains limited. Existing methods often rely on radiomic features, overlooking raw imaging and heterogeneous clinical data. We propose a deep learning (DL) framework that integrates multisequence magnetic resonance (MRI) and clinical variables for survival prediction. Our study compares a foundation model (FM) paired with a classifier against conventional convolutional neural networks (CNNs) to develop a robust model.

1 Introduction

Brain metastases are associated with a poor prognosis and high morbidity and mortality among cancer patients. Surgical intervention is often considered for selected patients, but it does not always improve quality of life or substantially prolong survival due to postoperative complications and the advanced nature of the disease [1, 2].

Deep learning (DL) has shown remarkable potential in radiology, enabling automated segmentation, diagnosis, and prediction of clinical outcomes [3]. Current studies on survival prediction for brain metastases have mainly focused on radiomic features, often omitting the direct use of imaging data [4, 5]. While several authors have highlighted the benefits of using multimodal data [6, 7], studies that directly integrate MRI images with more diverse clinical data—including previous treatments—and histological information remain limited [5].

To address these challenges, we leverage DL to integrate MRI sequences with clinical variables for survival prediction and individualized treatment guidance. We compare FMs, large neural networks trained on diverse datasets that perform well even with limited labeled data [8, 9], with traditional

CNNs. By combining multisequence MRI and clinical information, our approach aims not only to improve prognostic reliability but also to systematically evaluate and compare these two methodological frameworks in developing a clinically relevant decision-support model.

2 Data

We are collecting a dataset of 288 patients from Santa Maria Hospital, Lisbon, diagnosed with brain metastases between [period 2009-2024]. MRI scans were acquired 1–3 days before brain surgery, and include T1, T1 post-contrast (T1*), T2, and FLAIR sequences for all patients, as well as additional imaging modalities such as Apparent Diffusion Coefficient maps, perfusion when available. Moreover, it contains detailed clinical information, including tumor histology, primary tumor characteristics, and comprehensive treatment records.

2.1 Imaging Preprocessing

MRI sequences are first converted from DICOM to NIFTI format using `dcm2nii` [10] and preprocessed following the pipeline (`CBICA-CaPTk`). Briefly, MRI scans are first reoriented to LPS/RAI and then registered to the SRI24 atlas [11], including a temporary N4 bias field correction to compute the transformation matrix. Images are resampled to isotropic $1 \times 1 \times 1$ mm³ voxel spacing. Rigid registration is applied for T1, T2, and FLAIR to T1*, followed by rigid registration of T1* to the atlas. The resulting transformations are applied to all reoriented images, so that T1, T2, and FLAIR are mapped first to T1* space and then to atlas space. Finally, skull-stripping is performed using the HD-BET deep learning model [12]. For modeling, the MRI scans are stacked into 4D (4-channel) volumes in the order T1, T1*, T2, and FLAIR, which are then resized to $96 \times 96 \times 96$ voxels and intensity-normalized.

*Corresponding Author.

3 Preliminary Experiments

We first conducted an analysis of our target variable, survival, and framed the survival prediction task as a classification problem, taking into account both target variable distribution and clinical usability in the context of surgical decision-making. We examined the distribution of the target variable, checked for class balance, and evaluated correlations with follow-up time and mortality.

After applying the necessary preprocessing steps described previously, and with the goal of comparing the performance of FM and CNNs, we selected a Vision Transformer-based autoencoder implemented in PyTorch and MONAI. Specifically, we employed the ViTClassifier_block16 (ViT-16, with a dropout block size of $16 \times 16 \times 16$). This model leverages pretrained weights from large-scale medical imaging datasets [13]. Latent representations were computed during the forward pass for 151 patients¹. Each MRI sequence was initially evaluated independently, with the remaining sequences set to NULL, and missing latent variables treated as constants in the UMAP representation. Also, latent representations were computed considering the complete 4D volumes.

4 Preliminary Results

4.1 Survival as a Categorical Variable

To ensure clinically meaningful and balanced data, we stratified survival into three categories: <6 months, 6–21 months, and >21 months. This resulted in 21.87%, 26.39%, and 51.7% of patients in the short, medium, and long survival groups, respectively, with only 10 patients having medium survival and a follow-up shorter than 21 months (Figure 1).

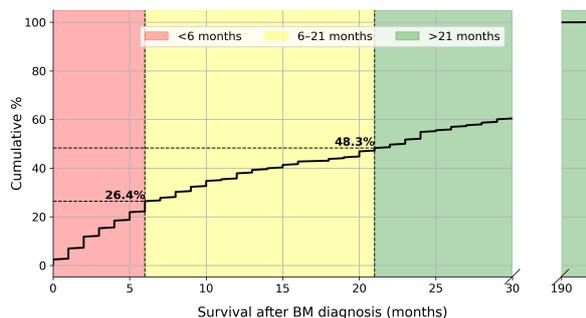


Figure 1. Cumulative distribution function (CDF) of survival after BM diagnosis, highlighting the three survival categories.

4.2 FM - Latent Embeddings

After processing the MRI sequences, as previously described, we computed and then examined how the

¹Of 288 patients, 151 had complete MRI data for analysis.

latent space distributions relate to the categorical survival variable, as visualized using UMAP (Figure 2). No clear separation was observed in the UMAP projection of the latent space across MRI sequences (Figure 2(b)). Although some local distinctions are visible, the latent representations do not exhibit a consistent or meaningful relationship with the predefined survival categories.

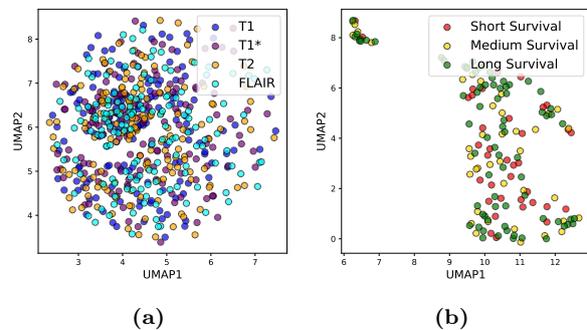


Figure 2. UMAP projection of latent representations from the pretrained ViT-16. (A) Colored by MRI sequence. (B) Colored by categorical survival groups.

5 Future Experiments

The current results suggest that the pre-trained FM embeddings do not accurately distinguish the defined categorical targets. Fine-tuning these embeddings, along with adapting the simple classification architecture proposed [13]—especially the average pooling and FC layers may be needed.

Future experiments will use our dataset with an 80/10/10 split for training, validation, and testing, along with external validation on a public dataset [14].

We will compare the FM with several architectures available in MONAI, including DenseNet121, ResNet18, and SENet154 (previously used by the authors [13]), as well as ResNet34/50/101, DenseNet169/201, and EfficientNet, applying Med3D transfer learning [15] where appropriate.

Our primary focus will be survival prediction, while also exploring outcomes such as primary tumor origin, BM relapse, molecular subtype, and markers, either via multi-output models or separate models per task.

Finally, we will develop an integrated multimodal framework combining MRI with clinical and histology data to enhance predictive performance and enable comprehensive multimodal analysis.

146 **References**

- 147 [1] D. N. Agney, A. M. Martin, P. J. Catalano, 201
148 A. J. Redig, N. U. Lin, E. Q. Lee, P. Y. Wen, 202
149 I. F. Dunn, W. L. Bi, S. E. Weiss, D. A. Haas- 203
150 Kogan, B. M. Alexander, and A. A. Aizer. 204
151 “Incidence and prognosis of patients with brain 205
152 metastases at diagnosis of systemic malignancy: a population-based study”. In: *Neuro-* 206
153 *Oncology* 19.11 (Oct. 2017), pp. 1511–1521. 207
154 DOI: [10.1093/neuonc/nox077](https://doi.org/10.1093/neuonc/nox077). 208
- 156 [2] S. Fadel and A. Ellaithy. “Brain metastasis: In- 209
157 cidence, trend analysis, and impact on survival 210
158 using SEER database (2010–2020)”. In: *Journal of Clinical Oncology* 43.16_suppl (2025), 211
159 pp. 2021–2021. DOI: [10.1200/JCO.2025.43.16\textunderscoresuppl.2021](https://doi.org/10.1200/JCO.2025.43.16.textunderscoresuppl.2021). eprint: 212
160 [https://ascopubs.org/doi/pdf/10.1200/](https://ascopubs.org/doi/pdf/10.1200/JCO.2025.43.16_suppl.2021) 213
161 [JCO.2025.43.16_suppl.2021](https://ascopubs.org/doi/pdf/10.1200/JCO.2025.43.16_suppl.2021). URL: [https://ascopubs.org/doi/abs/10.1200/JCO.](https://ascopubs.org/doi/abs/10.1200/JCO.2025.43.16%5C_suppl.2021) 214
162 [2025.43.16%5C_suppl.2021](https://ascopubs.org/doi/abs/10.1200/JCO.2025.43.16%5C_suppl.2021). 215
- 166 [3] M. A. Mazurowski, M. Buda, A. Saha, and 216
167 M. R. Bashir. “Deep learning in radiology: 217
168 An overview of the concepts and a survey of 218
169 the state of the art with focus on MRI”. In: *Journal of Magnetic Resonance Imaging* 49.4 219
170 (Apr. 2019). Epub 2018 Dec 21, pp. 939–954. 220
171 DOI: [10.1002/jmri.26534](https://doi.org/10.1002/jmri.26534). 221
- 173 [4] T. Öznacar, İ. P. Aral, H. Y. Zengin, and Y. 222
174 Tezcan. “Survival Prediction in Brain Metas- 223
175 tasis Patients Treated with Stereotactic Ra- 224
176 diosurgery: A Hybrid Machine Learning Ap- 225
177 proach”. In: *Brain Sciences* 15.3 (2025). ISSN: 226
178 2076-3425. DOI: [10.3390/brainsci15030266](https://doi.org/10.3390/brainsci15030266). 227
179 URL: [https://www.mdpi.com/2076-3425/](https://www.mdpi.com/2076-3425/15/3/266) 228
180 [15/3/266](https://www.mdpi.com/2076-3425/15/3/266). 229
- 181 [5] C. C. Volovăț, C. G. Buzea, D.-I. Boboc, M.-R. 230
182 Ostafe, M. Agop, L. Ochiuz, Ș. L. Burlea, 231
183 D. I. Rusu, L. Bujor, D. T. Iancu, and S. R. 232
184 Volovăț. “Hybrid Deep Learning for Survival 233
185 Prediction in Brain Metastases Using Multi- 234
186 modal MRI and Clinical Data”. In: *Diagnostics* 15.10 (2025). ISSN: 2075-4418. DOI: [10.](https://doi.org/10.3390/diagnostics15101242) 235
187 [3390/diagnostics15101242](https://doi.org/10.3390/diagnostics15101242). URL: [https://](https://www.mdpi.com/2075-4418/15/10/1242) 236
188 www.mdpi.com/2075-4418/15/10/1242. 237
- 190 [6] M. Islam, V. S. Vibashan, V. J. M. Jose, N. 238
191 Wijethilake, U. Utkarsh, and H. Ren. “Brain 239
192 Tumor Segmentation and Survival Prediction 240
193 Using 3D Attention U-Net”. In: *BrainLesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019*. Ed. by 241
194 A. Crimi and S. Bakas. Vol. 11992. Lecture 242
195 Notes in Computer Science. Springer, Cham, 243
196 2020. DOI: [10.1007/978-3-030-46640-4_25](https://doi.org/10.1007/978-3-030-46640-4_25). 244
197 URL: [https://doi.org/10.1007/978-3-](https://doi.org/10.1007/978-3-030-46640-4_25) 245
198 [030-46640-4_25](https://doi.org/10.1007/978-3-030-46640-4_25). 246
- 199 [7] T. Zhou, S. Ruan, S. Canu, P. Vera, and S. 247
200 Ruan. “Multi-modal Multi-channel Network 248
201 for Overall Survival Time Prediction of Brain 249
202 Tumor Patients”. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. Ed. by A. L. Martel et al. 250
203 Vol. 12262. Lecture Notes in Computer Science. 251
204 Springer, Cham, 2020. DOI: [10.1007/978-3-](https://doi.org/10.1007/978-3-030-59713-9_22) 252
205 [030-59713-9_22](https://doi.org/10.1007/978-3-030-59713-9_22). URL: [https://doi.org/](https://doi.org/10.1007/978-3-030-59713-9_22) 253
206 [10.1007/978-3-030-59713-9_22](https://doi.org/10.1007/978-3-030-59713-9_22). 254
- 207 [8] S. Ghamizi, G. Kanli, Y. Deng, M. Perquin, 255
208 and O. Keunen. *Brain Imaging Foundation 256
209 Models, Are We There Yet? A Systematic Re- 257
210 view of Foundation Models for Brain Imaging 258
211 and Biomedical Research*. 2025. arXiv: 2506. 259
212 13306 [eess.IV]. URL: [https://arxiv.org/](https://arxiv.org/abs/2506.13306) 260
213 [abs/2506.13306](https://arxiv.org/abs/2506.13306). 261
- 214 [9] V. van Veldhuizen, V. Botha, C. Lu, M. E. 262
215 Cesur, K. G. Lipman, E. D. de Jong, H. Hor- 263
216 lings, C. I. Sanchez, C. G. M. Snoek, L. Wes- 264
217 sels, R. Mann, E. Marcus, and J. Teuwen. 265
218 *Foundation Models in Medical Imaging – A 266
219 Review and Outlook*. 2025. arXiv: 2506.09095 267
220 [eess.IV]. URL: [https://arxiv.org/abs/](https://arxiv.org/abs/2506.09095) 268
221 [2506.09095](https://arxiv.org/abs/2506.09095). 269
- 222 [10] X. Li, P. S. Morgan, J. Ashburner, J. Smith, 270
223 and C. Rorden. “The first step for neuroimag- 271
224 ing data analysis: DICOM to NIfTI conver- 272
225 sion”. In: *Journal of Neuroscience Methods* 264 (2016), pp. 47–56. DOI: [10.1016/j.](https://doi.org/10.1016/j.jneumeth.2016.03.001) 273
226 [jneumeth.2016.03.001](https://doi.org/10.1016/j.jneumeth.2016.03.001). 274
- 227 [11] T. Rohlfing, N. M. Zahr, E. V. Sullivan, and A. 275
228 Pfefferbaum. “The SRI24 multichannel atlas 276
229 of normal adult human brain anatomy”. In: 277
230 *Human Brain Mapping* 31.5 (2010), pp. 798– 278
231 819. DOI: [10.1002/hbm.20906](https://doi.org/10.1002/hbm.20906). 279
- 232 [12] F. Isensee, M. Schell, I. Tursunova, G. Brug- 280
233 nara, D. Bonekamp, U. Neuberger, A. Wick, 281
234 H.-P. Schlemmer, S. Heiland, W. Wick, M. 282
235 Bendszus, K. H. Maier-Hein, and P. Kickinger- 283
236 eder. “Automated brain extraction of multi- 284
237 sequence MRI using artificial neural networks”. 285
238 In: *Human Brain Mapping* (2019), pp. 1–13. 286
239 DOI: [10.1002/hbm.24750](https://doi.org/10.1002/hbm.24750). 287
- 240 [13] M. Chen, M. Zhang, L. Yin, and et al. “Med- 288
241 ical image foundation models in assisting di- 289
242 agnosis of brain tumors: a pilot study”. In: 290
243 *European Radiology* 34 (2024), pp. 6667–6679. 291
244 DOI: [10.1007/s00330-024-10728-1](https://doi.org/10.1007/s00330-024-10728-1). 292
- 245 [14] D. Ramakrishnan, L. Jekel, S. Chadha, A. 293
246 Janas, H. Moy, N. Maleki, M. Sala, M. Kaur, 294
247 G. C. Petersen, S. Merkaj, M. von Reppert, 295
248 U. Baid, S. Bakas, C. Kirsch, M. Davis, K. 296
249 Bousabarah, W. Holler, M. Lin, M. Wester- 297
250 hoff, S. Aneja, F. Memon, and M. Aboian. *A 298
251 Large Open Access Dataset of Brain Metas-* 299
252 *253
254
255
256*

- 257 *tasis 3D Segmentations on MRI with Clin-*
258 *ical and Imaging Feature Information.* The
259 Cancer Imaging Archive, Version 1 [dataset].
260 2023. DOI: [10.7937/6be1-r748](https://doi.org/10.7937/6be1-r748). URL: [https:](https://doi.org/10.7937/6be1-r748)
261 [//doi.org/10.7937/6be1-r748](https://doi.org/10.7937/6be1-r748).
- 262 [15] S. Chen, K. Ma, and Y. Zheng. “Med3D:
263 Transfer Learning for 3D Medical Image Anal-
264 ysis”. In: *arXiv preprint arXiv:1904.00625*
265 (2019).