Journal of Neuro-Oncology
https://doi.org/10.1007/s11060-024-04630-5

RESEARCH

Comparative analysis of the spatial distribution of brain metastases across several primary cancers using machine learning and deep learning models

Saeedeh Mahmoodifar[1] · Dhiraj J. Pangal[2] · Josh Neman[2] · Gabriel Zada[2] · Jeremy Mason[3,4] · Bodour Salhia[5] · Tehila Kaisman-Elbaz[6] · Selcuk Peker[7] · Yavuz Samanci[7] · Andréanne Hamel[8] · David Mathieu[8] · Manjul Tripathi[9] · Jason Sheehan[10] · Stylianos Pikis[10] · Georgios Mantziaris[10] · Paul K. Newton[11]

Received: 27 February 2024 / Accepted: 2 March 2024

Abstract

Objective Brain metastases (BM) are associated with poor prognosis and increased mortality rates, making them a significant clinical challenge. Studying BMs can aid in improving early detection and monitoring. Systematic comparisons of anatomical distributions of BM from different primary cancers, however, remain largely unavailable.

Methods To test the hypothesis that anatomical BM distributions differ based on primary cancer type, we analyze the spatial coordinates of BMs for five different primary cancer types along principal component (PC) axes. The dataset includes 3949 intracranial metastases, labeled by primary cancer types and with six features. We employ PC coordinates to highlight the distinctions between various cancer types. We utilized different Machine Learning (ML) algorithms (RF, SVM, TabNet DL) models to establish the relationship between primary cancer diagnosis, spatial coordinates of BMs, age, and target volume.

Results Our findings revealed that PC1 aligns most with the Y axis, followed by the Z axis, and has minimal correlation with the X axis. Based on PC1 versus PC2 plots, we identified notable differences in anatomical spreading patterns between Breast and Lung cancer, as well as Breast and Renal cancer. In contrast, Renal and Lung cancer, as well as Lung and Melanoma, showed similar patterns. Our ML and DL results demonstrated high accuracy in distinguishing BM distribution for different primary cancers, with the SVM algorithm achieving 97% accuracy using a polynomial kernel and TabNet achieving 96%. The RF algorithm ranked PC1 as the most important discriminating feature.

Conclusions In summary, our results support accurate multiclass ML classification regarding brain metastases distribution.

Keywords Brain metastases · Principal components · Deep learning models · Pan cancer analysis

Introduction

It is well established that different primary cancer types, and different molecular subtypes distribute metastases preferentially to different locations [1–8], although a quantitative understanding of the spatial distribution of metastatic disease, and the temporal ordering of when these metastases first appear at the different locations remains far less understood [3]. Although metastases to the brain are not usually the location of the first metastatic site for any primary cancer type [4], the presence of BMs portend poor prognosis for the patient, regardless of cancer subtype. The advancements in treatment regimens, including the development of immunologic therapies have increased life expectancies for a number of primary cancers, and brought new importance to the study of BMs, their natural progression and causes for growth.

Recently, progress has been made in quantifying the spatial distribution of brain metastases for breast cancer patients of different molecular subtypes [9], showing quantitively distinct patterns in some categories. The underlying hypothesis rests on the notion that different cancers require different environments for growth [10, 11], and therefore are more or less likely to metastasize in certain regions of the brain. We aim to expand on the work performed by our group in prior studies [1–9], by exploring the predictive ability of a machine learning model to determine the primary subtype of cancer given spatial information about its three-dimensional location, as well as age at treatment and target volume. The

✉ Paul K. Newton
newton@usc.edu

Extended author information available on the last page of the article

Published online: 02 April 2024

potential ability for machine learning models to accurately identify the primary cancer type from these small set of features would indicate that these differences are distinct enough to be discerned, which might further motivate the search for underlying biological explanations for these differences.

We demonstrate that using spatial data as the primary means of input, a machine learning model can accurately parse out the primary cancer subtype from a large dataset of brain metastases from a national brain tumor metastasis registry.

Methods

Dataset

Data used in this analysis is taken from the International Radiosurgery Registry Foundation (IRRF) and all patients underwent gamma-knife radiosurgery (GKRS) for the treatment of brain metastases which are labeled based on the primary cancer types Breast, Lung, Melanoma, Renal, and Colon. The dataset consists of six features including sex, age, target volume, and stereotactic Cartesian coordinates X, Y, and Z of a total of 3949 intracranial metastases. See the data summarized in Tables S1 and S2.

Principal component analysis (PCA)

The principal component coordinates are a data driven orthogonal coordinate system intended to highlight the directions of the greatest spread of the data, with PC1 as the direction of the largest variance and PC2 and PC3 as the directions that capture the remaining variations orthogonal to the first principal component and to each other. PCA is used to identify patterns in a dataset and as a method of dimensionality reduction for high dimensional datasets by identifying new uncorrelated features which allows better visualization of the dataset [12]. We use PCA from the Scikit-learn library in Python for our analysis [13].

Synthetic minority over-sampling technique (SMOTE)

SMOTE is a data augmentation method used to address a class imbalance in supervised machine learning problems. Class imbalance occurs when one class of a classification problem has significantly fewer samples than the other classes, which can lead to poor performance of the classifier on the minority class. SMOTE creates synthetic samples of the minority class by interpolating between existing minority class samples. The method selects a minority class sample and identifies its k nearest neighbors in the feature space. SMOTE then creates a new sample by randomly selecting one of the ”k” nearest neighbors and creating a synthetic sample between the original sample and each of its neighbors that is a linear combination of the original and selected neighbors. The process repeats until the desired balance between the classes is achieved. The synthetic samples created by SMOTE increase the size of the minority class, making it more representative and improving the classifier’s ability to learn the patterns in the minority class [14].

Dataset preprocessing

PCA is sensitive to the scaling of the variables in the dataset. Variables that have larger magnitudes will dominate the variance and may obscure the contribution of other variables that have smaller magnitudes. Scaling the variables to a common scale ensures that all variables are equally important in the analysis. Different variables in the dataset have different units of measurement, and these units can affect the calculation of the principal components. Scaling the variables to unit variance (i.e., standardizing) removes the units of measurement and allows the components to be calculated based on the correlations between the variables. In order to ensure that the result of the PCA is representative of the underlying patterns in the data, we scale our features before performing the PCA. We use the StandardScaler from the Scikit-learn package in Python which set the mean to zero and the variance to one [13]. In order to reduce the effect of class imbalance in the dataset, we use the Synthetic Minority Over-sampling Technique using the SMOTE from the imbalanced-learn library in Python [15]. We split up the dataset into 90% training and 10% testing. The reported evaluation metrics in Table 1 correspond to the testing dataset.

Random forest (RF)

Random Forest is a supervised machine learning algorithm for classification and regression tasks. It is an ensemble learning method that combines multiple decision trees to make a final prediction. During the training process, Random Forest builds a large number of decision trees by using a randomly selected portion of the training data along with a randomly selected subset of the available features. Each tree is built

Table 1 Evaluation metrics for different machine learning and deep learning models

Model	Accuracy	Precision	Recall	F1-score
RF	89%	89%	89%	89%
SVM-linear	71%	78%	71%	72%
SVM-poly	97%	97%	97%	97%
SVM-rbf	94%	94%	94%	94%
TabNet	96%	97%	96%	96%

independently, and at each split, the algorithm selects the best feature to split on among a random subset of features. This randomness helps reduce overfitting and improves the model's generalization performance. Once the trees are built, the Random Forest algorithm combines their predictions to make a final prediction. In classification tasks, the class with the most votes is selected, and in regression tasks, the mean or median of the individual tree predictions is taken [16]. A random forest classifier is built based on RandomForestClassifier from the Scikit-learn package [13].

Support vector machine (SVM) and one v. all (OvA)

The linear SVM algorithm aims to find a hyperplane that separates two tumor classes to maximize the distance between the hyperplane and the nearest samples from each class. In order to determine the maximum separation distance between classes, the dot products of support vectors and the classes must be computed [17]. The main concept behind this is identifying the largest margin between the classes. In cases where the data is not linearly separable, SVMs can use a kernel function to transform the data into a higher dimensional space that a hyperplane can separate. The kernel functions used in this study are linear, polynomial, and radial basis functions. For transitioning from binary to multiclass classification, we adopt a One-vs-All (OvA) approach [18]. The OvA strategy involves training multiple binary classifiers, each distinguishing one class from all the others. For each class, a binary classifier is trained to distinguish between that class and all the other classes combined. This results in a set of binary classifiers, one for each class. The classifier with the highest confidence score is selected as the predicted class during prediction. Kernel SVMs are helpful when the underlying relationships in the data are non-linear and can model complex non-linear decision boundaries by transforming the input features into a higher-dimensional space. This study uses three different kernels: Radial Basis Function, Polynomial, and Linear kernel. For our analysis, we utilize the SVC algorithm from the Scikit-learn library in Python [13].

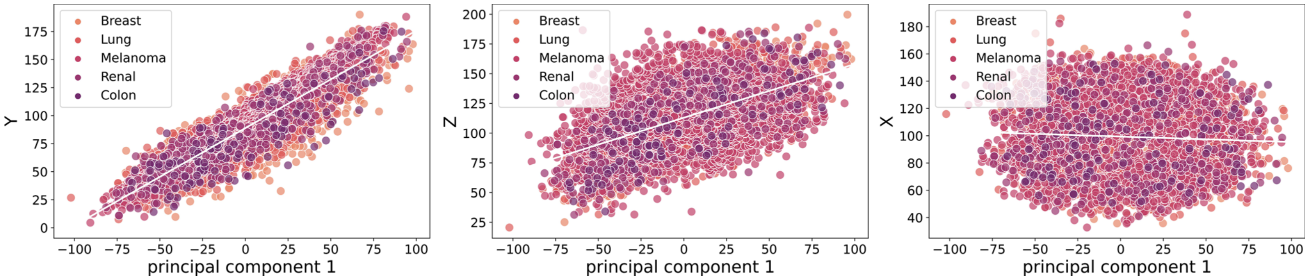
TabNet

Deep learning algorithms have generally been successful in classifying images or audio but not tabular data [19]. TabNet, a deep neural network (DNN) tailored for learning from tabular data, employs a distinctive architecture known as the TabNet encoder [20]. In this architecture, sequential multisteps (Nsteps) are a pivotal component. Each step, denoted as i, leverages processed information from the previous step $(i-1)$ to make decisions regarding feature utilization. These decisions culminate in processed feature representations, which, in turn, play a critical role in the overarching decision-making process.

Results

In Fig. 1 we show 2D scatter plots of BM locations for five different primary cancer types (Breast, Lung, Melanoma, Renal, and Colon) plotted with their PC1 component versus each of the (X, Y, Z) Cartesian coordinates in 3D space. Figure 1a shows that PC1 correlates most strongly with the Y coordinate (front-to-back), next, Fig. 1b shows the correlation with the Z coordinate (top-to-bottom), while Fig. 1c shows there is very little linear correlation with the X coordinate (side-to-side). See [9] Fig. S1 for a more detailed plot of the coordinate systems used with the GKRS stereotactic headset which measures BM locations.

Our conclusion from these comparisons is that the (PC1, PC2) plane offers an optimal [12] reduced dimension plane that most accurately will depict the differences in the spatial distributions of BMs for the five different cancer types, in addition to the other features from the data set. The side-to-side X coordinate distribution is the least important of the



(a) (b) (c)

Fig. 1 Brain metastases scatter plots for five different primary cancers along Principal Component 1 axis (PC1) vs. X, Y, and Z coordinates showing strongest linear correlation between Y axis and PC1 axis. a) 2D projection of data onto (PC1, Y) plane and linear curve fit; b) 2D projection of data onto (PC1, Z) plane and linear curve fit; c) 2D projection of data onto (PC1, X) and linear curve fit

three, reflecting the fact that the five cancers all distribute their BMs more or less symmetrically across the midline. Given that it is the (Y, Z) coordinate plane that mostly captures the important differences in Cartesian BM locations, we show in Fig. 2 PC1 for each of the primary cancer types projected onto this plane. Both the means (the base of each coordinate arrow) and the directions of each PC1 vector are different for each primary cancer type as can easily be seen. Note the similarity, however, between the direction of the PC1 axis associated with lung and renal cancers, with only the mean basepoint shifting between the two.

In Fig. 3 we focus on depicting the BM locations in the PC1 vs. PC2 planes (the optimal reduced-order plane). In Figs. 3a, 3b we show the distributions of the two cancer types that are most distinct with respect to their spatial distributions: Breast vs. Lung (3a), and Breast vs. Renal (3b). We indicate these differences by plotting the linear curve fits to each of the cancer types on the same plots, showing both their means and orientation of the linear curves are very distinct. By contrast, Figs. 3c, 3d show the distributions that are most similar: Lung vs. Melanoma (3c), and Lung vs. Renal (3d). Note the similarity of their means and linear curve fits as compared with those in Figs. 3a, 3b. The linear curve fits in these four plots are not intended to indicate that the data closely follows a linear regression model, but only meant to show the most apparent differences/similarities in the spread of points along the regression line (i.e a useful visual guide).

We use three different Machine Learning, and Deep Learning algorithms: Random Forest model, Support Vector Machine (SVM) and TabNet to see how well each can distinguish between the BM spatial distributions associated with the five primary cancer types. These were chosen based on their ease of implementation for our dataset and our assessment of their high liklihood of distinguishing subtle pattern differences in our data subgroups. See discussions of these algorithms in the Methods section and references therein.

The first important observation is shown in Fig. 4 where we plot the relative importance of the top 8 most important features from the data. PC1 is identified as the most important feature, followed by the Z coordinate, the Y coordinate, then PC2, PC3, followed by Age at treatment, X coordinate, and Target volume. Taken together, our conclusion is consistent with our previous observations, that (PC1, PC2) are a more efficient coordinate system to use than (Y,Z) given that the PC1 direction captures most of the spread in the (Y,Z) plane. In addition, Age at treatment seems to be a more important variable than Target volume in distinguishing spatial BM distributions. Table 1 summarizes key metrics (Accuracy, Precision, Recall, F1-score) associated with the Random Forest (RF) method, and three different Support Vector Machine (SVM) methods: SVM-linear; SVM-poly; and SVM-rbf as well as TabNet. With all metrics, the SVM-poly method performs best, scoring at 97% on the test data in each category. In Table 1, Precision is the number of correctly identified members of a class divided by the number of times the model predicted that class; Recall is the number of members of a class that the classifier identified correctly divided by the total number of members in that class; and F1-score is a combination of Precision and Recall combined

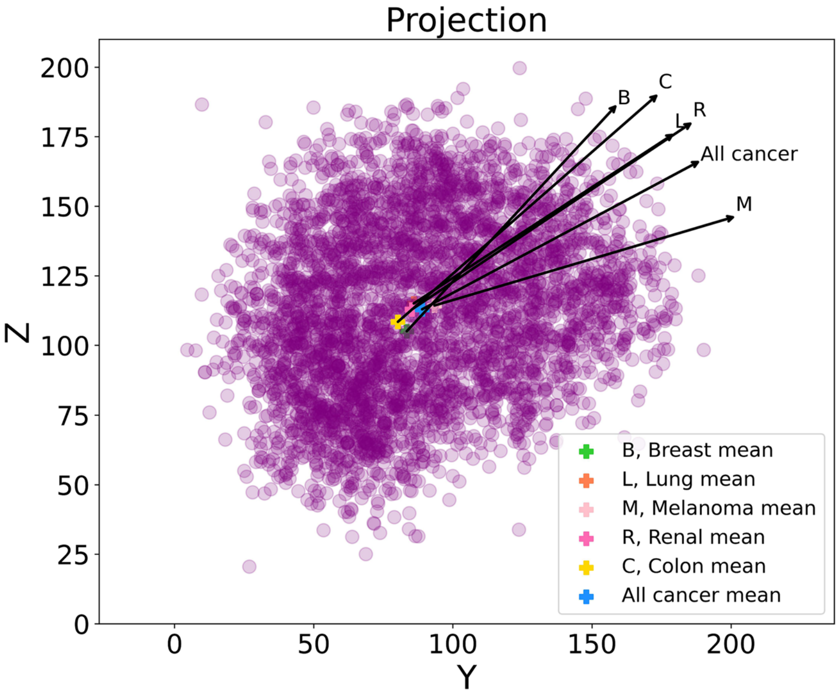


Fig. 2 2D projection of scatter plot of all cancer metastatic brain tumors onto (Y, Z) plane showing the Principal component 1 axis for each cancer type separately and with respect to all cancer types together. Violet crosses indicate the means of each cancer and the yellow cross indicates the mean of all data points

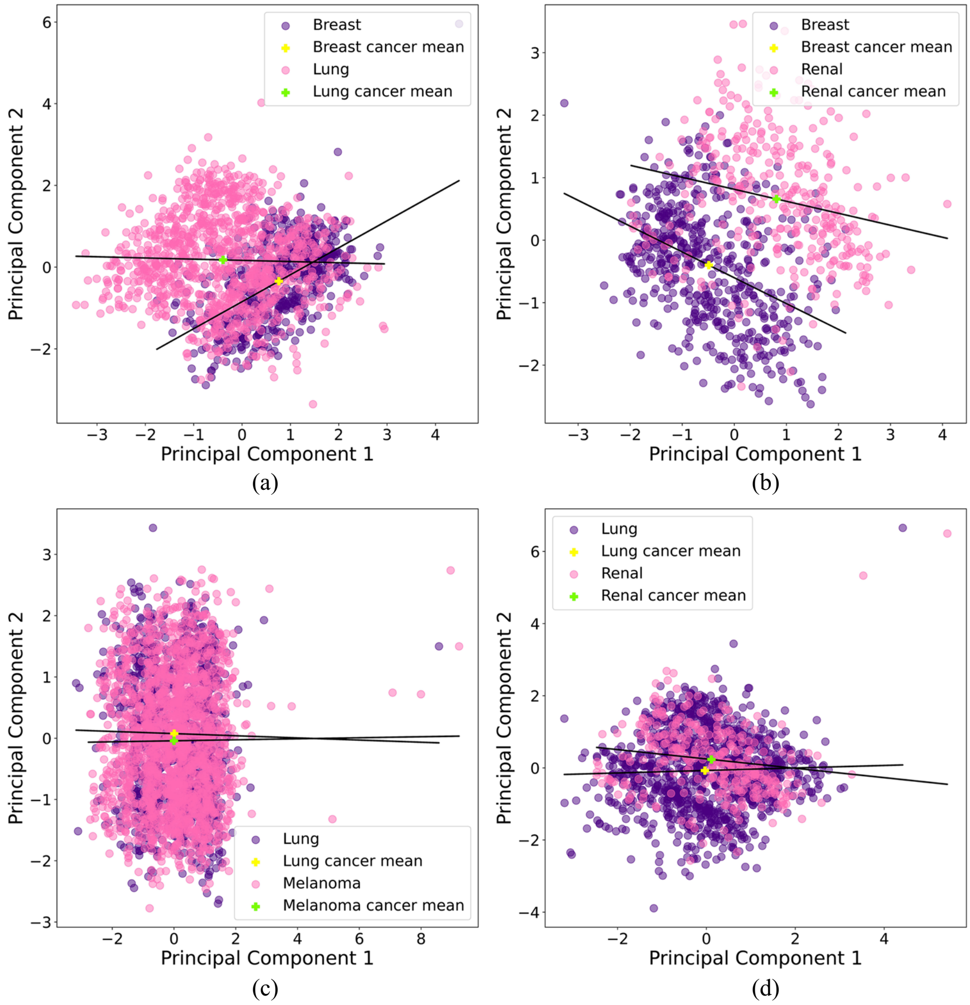


Fig. 3 Scatter plot of pair cancer types onto (PC1, PC2) axes. The black line indicates the linear curve fit is not meant to imply that the data is spread linearly, but is useful to draw attention to differences in the two data sets being compared. Yellow and green crosses show the means. (a) and (b) plots have the most distinct brain metastasis distributions (lung vs. breast cancers, and breast vs. renal cancers); (c) and (d) have the most similar brain metastasis distributions (lung vs. melanoma cancers, and lung vs. renal cancers)

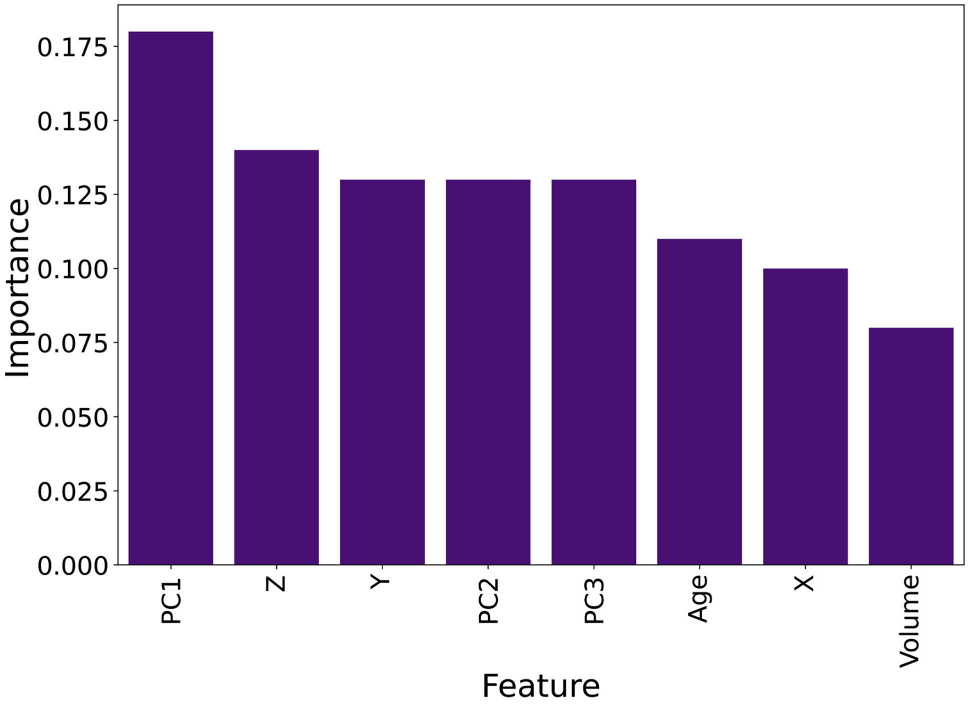


Fig. 4 Bar plot showing the feature importance of the Random Forest model. (PC1,PC2,PC3) coordinate features are collectively more important than (X,Y,Z) coordinate features. Age at Treatment is a more important non-coordinate feature than Target volume

into one single metric. By contrast, SVM-linear performs the poorest in each category.

Discussion

The primary focus of this study was to analyze the spatial distribution of BMs across diverse primary cancers and evaluate the capacity of machine learning models to distinguish between them. The results underscore the significance of the first principal component (PC1), which exhibited substantial alignment with the Y and Z axes (Figs. 1a, 1b), while the X axis showed minimal correlation (Fig. 1c), consistent with left-right symmetries across primary cancer types. Utilizing the (PC1, PC2) plane as an optimally reduced dimension plane proved superior in depicting differences in BM spatial distributions compared to the (Y, Z) coordinate plane. Notably, distinctive spatial distribution patterns were observed for Breast vs. Lung and Breast vs. Renal cancers [21], emphasizing potential specificity to primary cancer types. Conversely, similarities were noted in distributions for Lung vs. Melanoma and Lung vs. Renal cancers, hinting at shared mechanisms in their development. Machine learning algorithms, including Random Forest (RF), Support Vector Machine (SVM), and TabNet, effectively differentiated between primary cancer subtypes based on BM spatial distributions. PC1 emerged as a pivotal feature, with the SVM algorithm using a polynomial kernel achieving a notable 97% accuracy, showcasing robust differentiation capabilities. The high accuracy achieved by these models in most cases not only suggests the presence of distinct differences in the spatial distribution of BMs across primary cancer types but also indicates that the translation of these distributions onto the first principal component (PC1) further enhances the differentiation capabilities as indicated by its standing as the most important feature in the RF algorithm. This observation implies that utilizing the PC1, which already highlights differences in spatial distribution, can be a robust approach for parsing out these distinctions among primary cancer subtypes and should be an important component in using ML methods on larger data sets. The downstream effects of developing ML and DL models for BM subtyping could be multifold. First, it is estimated that no primary tumor source is identified in roughly 15% of patients, even after imaging workup [1]. Even for patients in which BM have known primary cancer diagnosis, there are often instances where neurologic symptoms and brain MRI are the first scans which demonstrate tumor burden. A high- fidelity test could at the minimum, key in radiologists and oncologists to look out for a particular subtype or unusual patterns. We also note the clinical utility in shrinking the radiation field based on knowledge of tumor-specific distributions to minimize unwanted consequences of high radiation doses. Additionally, by addressing phenotypic, tumoral behavior characteristics (e.g. where it metastasizes), and exploring molecular traits which have overlap irrespective of primary cancer subtype (i.e. where it came from), we may unlock new options for therapeutic targets that are shared between seemingly disparate cancer subtypes. Although this study focuses on the development of data analytics tools, we mention some of the important biological hypotheses for why brain metastases from different tumor types could favor one region of the brain over another. First is the ’seed-and-soil’ hypothesis developed over 100 years ago (discussed in the contex of brain metastases in [10]) which postulates that cells from the primary tumor (’seed’) metastasize to specific areas of the brain (’soil’) due to its unique microenvironment that both attracts and allows the tumor to grow [11]. Both the molecular and phenotypic ’compatibility’ of the seed and soil are important in this framework and have been explored [11]. A second hypothesis is differening vascularization patterns in the brain [9] and the associated cellular communication between tumor cells, brain pericytes, astrocytes, and vascular endothelial cells partly responsible for growth and stimulation of vasculature required for spread. For example, as discussed in [1], it is known that certain primary cancers (melanoma, NSCLC, and breast cancers) can co-opt the vasculature and grow preferentially along existing vessels, while other lung cancers can instigate early angiogenesis by effecting differing cellular receptors and protein expression patterns specific to tumor type.

Limitations and challenges

Limitations and challenges associated with this study include the fact that it is a retrospective study from a specific tumor registry. It would be desirable to carry out a prospective study that could highlight differences in treatment history (e.g. chemotherapy, radiation) and differences in brain metastasis distributions from different geographic or demographic populations.

Challenges include those associated with differences in brain size and shapes, and pinpointing the exact location of each BM (defined as the center) given their varying volumes and shapes using different imaging modalities at different cancer center facilities.

Conclusions

For the purposes of distinguishing the spatial distribution of brain metastases associated with the five primary cancer types under study, we find that the optimal data-designed coordinates PC1 vs. PC2, as opposed to the Cartesian Y-Z coordinate plane offers the best reduced dimensional projection in which to highlight differences in the spread of the BM data. As a variable in our feature-based machine learning approaches, PC1 emerges as the single most important feature to distinguish the spatial patterns. Instead of the (X,Y,Z) features in our ML approaches, the best set of features to use are (PC1, PC2, PC3), with Age at Treatment being more important than Target volume, but less important than the PC variables. The SVM-poly ML method performs very well (97% on test data by all metrics) in distinguishing among the five cancer types based on their BM distributions. We believe with more data, and better optimization of the ML and DL pipelines, ML and DL methods offer a very promising approach towards discerning potentially subtle differences in BM distributions associated with primary tumor type.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11060-024-04630-5.

Author Contributions S.M., D.P., J.M., and P.N. wrote the main manuscript text. D.P., B.S., J.N., and G.Z. provided and interpreted the data. S.M. and P.N. prepared figures and plots. T.K.E, S.P., Y.S., A.H., D.M., M.T., J.S., S.P., and G.M. collected the data. All authors reviewed the manuscript.

Funding Partial funding through the USC Norris Comprehensive Cancer Center's Multi-Level Cancer Risk Prediction Models pilot Project Award, 'Molecular, Clinical and Neuro-imaging Determinants of Spatiotemporal Pathogenesis of Cancer-Specific Brain Metastases: Data Analysis and Longitudinal Modeling' (12/01/2020-11/30/2021) is gratefully acknowledged.

Data Availability Data that are used in this study are part of the International Radiosurgery Registry Foundation (Study #: IRRF 02-15-2022), & Topography of Brain Metastases by Primary Cancer Genetic Subtypes & Data collection was approved by institutional review boards at each of the institutions affiliated with the IRRF. Due to retrospective design, informed consents were not obtained.

Declarations

Competing interests The authors declare no competing interests.

References

1. Cardinal T, Pangal D, Strickland BA, Newton P, Mahmoodifar S, Mason J, Craig D, Simon T, Tew BY, Yu M et al (2022) Anatomical and topographical variations in the distribution of brain metastases based on primary cancer origin and molecular subtypes: a systematic review. Neuro-Oncol Adv 4(1):170. https://doi.org/10.1093/noajnl/vdab170
2. In GK, Mason J, Lin S, Newton PK, Kuhn P, Nieva J (2017) Development of metastatic brain disease involves progression through lung metastases in egfr mutated non-small cell lung cancer. Converg Sci Phys Oncol 3(3):035002. https://doi.org/10.1088/2057-1739/aa7a8d
3. Newton PK, Mason J, Venkatappa N, Jochelson MS, Hurt B, Nieva J, Comen E, Norton L, Kuhn P (2015) Spatiotemporal progression of metastatic breast cancer: a markov chain model highlighting the role of early metastatic sites. NPJ Breast Cancer 1(1):1–9. https://doi.org/10.1038/npjbcancer.2015.18
4. Newton PK, Mason J, Hurt B, Bethel K, Bazhenova L, Nieva J, Kuhn P (2014) Entropy, complexity and markov diagrams for random walk cancer models. Sci Rep 4(1):7558. https://doi.org/10.1038/srep07558
5. Newton PK, Mason J, Bethel K, Bazhenova L, Nieva J, Norton L, Kuhn P (2013) Spreaders and sponges define metastasis in lung cancer: a markov chain monte carlo mathematical model. Can Res 73(9):2760–2769. https://doi.org/10.1158/0008-5472.CAN-12-4488
6. Newton PK, Mason J, Bethel K, Bazhenova LA, Nieva J, Kuhn P (2012) A stochastic markov chain model to describe lung cancer growth and metastasis. PLoS ONE 7(4):34637. https://doi.org/10.1371/journal.pone.0034637
7. Schroeder T, Bittrich P, Kuhne J, Noebel C, Leischner H, Fiehler J, Schroeder J, Schoen G, Gellisen S (2020) Mapping distribution of brain metastases: does the primary tumor matter? J Neuro-Oncol 147:229–235. https://doi.org/10.1007/s11060-020-03419-6
8. Neman J, Franklin M, Madaj Z, Deshpande K, Triche TJ, Sadlik G, Carmichael JD, Chang E, Yu C, Strickland BA et al (2021) Use of predictive spatial modeling to reveal that primary cancers have distinct central nervous system topography patterns of brain metastasis. J Neurosurg 136(1):88–96. https://doi.org/10.3171/2021.1.JNS203536
9. Mahmoodifar S, Pangal DJ, Cardinal T, Craig D, Simon T, Tew BY, Yang W, Chang E, Yu M, Neman J et al (2022) A quantitative characterization of the spatial distribution of brain metastases from breast cancer and respective molecular subtypes. J Neuro-Oncol 160(1):241–251. https://doi.org/10.1007/s11060-022-04147-9
10. Fidler IJ, Yano S, Zhang R-d, Fujimaki T, Bucana CD (2002) The seed and soil hypothesis: vascularisation and brain metastases. Lancet Oncol 3(1):53–57. https://doi.org/10.1016/s1470-2045(01)00622-2
11. Fidler IJ (2011) The role of the organ microenvironment in brain metastasis. In: Seminars in cancer biology, vol 21. Elsevier, pp 107–112. https://doi.org/10.1016/j.semcancer.2010.12.009
12. Kirby M (2000) Geometric data analysis: an empirical approach to dimensionality reduction and the study of patterns. John Wiley & Sons Inc, New York
13. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V et al (2011) Scikit-learn: machine learning in python. J Mach Learn Res 12:2825–2830
14. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) Smote: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357. https://doi.org/10.1613/jair.953
15. Lemaître G, Nogueira F, Aridas CK (2017) Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning. J Mach Learn Res 18(1):559–563
16. Breiman L (2001) Random forests. Mach Learn 45:5–32
17. Gupta P, Garg S (2020) Breast cancer prediction using varying parameters of machine learning models. Procedia Comput Sci 171:593–601. https://doi.org/10.1016/j.procs.2020.04.064
18. Ramaswamy S, Tamayo P, Rifkin R, Mukherjee S, Yeang C-H, Angelo M, Ladd C, Reich M, Latulippe E, Mesirov JP et al (2001) Multiclass cancer diagnosis using tumor gene expression signatures. Proc Natl Acad Sci 98(26):15149–15154. https://doi.org/10.1073/pnas.211566398

19. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778. https://doi.org/10.1109/CVPR.2016.90
20. Arik SÖ, Pfister T (2021) Tabnet: attentive interpretable tabular learning. In: Proceedings of the AAAI conference on artificial intelligence, vol 35, pp 6679–6687. https://doi.org/10.1609/aaai.v35i8.16826
21. Quattrocchi CC, Errante Y, Gaudino C, Mallio CA, Giona A, Santini D, Tonini G, Zobel BB (2012) Spatial brain distribution of intra-axial metastatic lesions in breast and lung cancer patients. J Neuro-Oncol 110:79–87. https://doi.org/10.1007/s11060-012-0937-x

Authors and Affiliations

Saeedeh Mahmoodifar[1] · Dhiraj J. Pangal[2] · Josh Neman[2] · Gabriel Zada[2] · Jeremy Mason[3,4] · Bodour Salhia[5] · Tehila Kaisman-Elbaz[6] · Selcuk Peker[7] · Yavuz Samanci[7] · Andréanne Hamel[8] · David Mathieu[8] · Manjul Tripathi[9] · Jason Sheehan[10] · Stylianos Pikis[10] · Georgios Mantziaris[10] · Paul K. Newton[11]

[1] Department of Physics & Astronomy, University of Southern California, Los Angeles, CA 90089, USA

[2] Department of Neurosurgery, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA

[3] Catherine & Joseph Aresty Department of Urology, Institute of Urology, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA

[4] Convergent Science Institute in Cancer, Michelson Center for Convergent Bioscience, University of Southern California, Los Angeles, CA 90089, USA

[5] Department of Translational Genomics Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA

[6] Rose Ella Burkhardt Brain Tumor & Neuro-Oncology Center, Neurological Institute, The Cleveland Clinic, Cleveland, OH 44195, USA

[7] Department of Neurosurgery, Koc University School of Medicine, Istanbul, Turkey

[8] Department of Neurosurgery, Université de Sherbrooke, Centre de recherche du CHUS, QC, Canada

[9] Department of Neurosurgery, Postgraduate Institute of Medical Education and Research, Chandigarh, India

[10] Department of Neurological Surgery, University of Virginia, Charlottesville, VA 22903, USA

[11] Department of Aerospace & Mechanical Engineering, Mathematics, Quantitative & Computational Biology, and Lawrence J. Ellison Institute for Transformative Medicine, University of Southern California, Los Angeles, CA 90089, USA