Integrative Deep Learning for PanCancer Molecular Subtype Classification Using Histopathological Images and RNAseq Data

Fatima Zare
Computer Science and Engineering Department, University of Connecticut
Storrs, Connecticut
fatemeh.zare@uconn.edu

Javad Noorbakhsh*
The Jackson Laboratory For Genomic Medicine
Farmington, Connecticut
jnoorbak@broadinstitute.org

Tianyu Wang
Computer Science and Engineering Department, University of Connecticut
Storrs, Connecticut
tianyu.wang@uconn.edu

Jeffrey H. Chuang
The Jackson Laboratory For Genomic Medicine,
UConn Health, Department of Genetics and Genome Sciences
Farmington, Connecticut
jeff.chuang@jax.org

Sheida Nabavi†
Computer Science and Engineering Department, University of Connecticut
Storrs, Connecticut
sheida.nabavi@uconn.edu

ABSTRACT

Deep learning has recently become a key methodology for the study and interpretation of cancer histology images. The ability of convolutional neural networks (CNNs) to automatically learn features from raw data without the need for pathologist expert knowledge, as well as the availability of annotated histopathology datasets, have contributed to a growing interest in deep learning applications to histopathology. In clinical practice for cancer, histopathological images have been commonly used for diagnosis, prognosis, and treatment. Recently, molecular subtype classification has gained significant attention for predicting standard chemotherapy's outcomes and creating personalized targeted cancer therapy. Genomic profiles, especially gene expression data, are mostly used for molecular subtyping. In this study, we developed a novel, PanCancer CNN model based on Google Inception V3 transfer learning to classify molecular subtypes using histopathological images. We used 22,484 Haemotoxylin and Eosin (H&E) slides from 32 cancer types provided by The Cancer Genome Atlas (TCGA) to train and evaluate the model. We showed that by employing deep learning, H&E slides can be used for classification of molecular subtypes of solid tumor samples with the high area under curves (AUCs) (micro-average= 0.90; macro-average=0.90). In cancer studies, combining histopathological images with genomic data has rarely been explored. We investigated the relationship between features extracted from H&E images and features extracted from gene expression profiles. We observed that the features from these two different modalities (H&E images and gene expression values) for molecular subtyping are highly correlated. We, therefore, developed an integrative deep learning model that combines histological images and gene expression profiles. We showed that the integrative model improves the overall performance of the molecular subtypes classification ((AUCs) micro-average= 0.99; macro-average=0.97). These results show that integrating H&E images and gene expression profiles can enhance accuracy of molecular subtype classification.

KEYWORDS

Deep Learning, PanCancer, Convolutional Neural Networks, Computational Histopathology, Digital Pathology, Histology Image Analysis, Transfer Learning

ACM Reference format:
Fatima Zare, Javad Noorbakhsh, Tianyu Wang, Jeffrey H. Chuang, and Sheida Nabavi. 2020. Integrative Deep Learning for PanCancer Molecular Subtype Classification Using Histopathological Images and RNAseq Data . In *Proceedings of Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, Virtual Event, USA, September 21–24, 2020 (BCB '20),* 8 pages.
https://doi.org/10.1145/3388440.3412414

1 INTRODUCTION

Histological images confer important information for the assessment of disease. Histology slides contain rich phenotypic information for cancer diagnosis and prognosis and are widely used by pathologists. However, cancer is a heterogeneous disease, and heterogeneity in cancer is not only limited to the differences among different patients but often occurs within a single patient. This inter-patient and intertumoral heterogeneity may pose major challenges, leading to inconsistencies in diagnosis [1–3, 11].

In addition to histopathologic images, molecular features such as genetic changes and signatures of gene expression are now

*The author is currently with Broad Institute of MIT and Harvard, Cambridge, Massachusetts.

†Corresponding author. JHC acknowledges support from NCI grant R01CA230031, PI: Chuang.

BCB '20, September 21–24, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-7964-9/20/09…$15.00
https://doi.org/10.1145/3388440.3412414

commonly used for the prediction of clinical cancer outcomes. After the advent of microarray analysis, many studies established a comprehensive molecular classification of cancer using genomic data. These studies have confirmed that cancers can be categorized into molecular subtypes based on distinct genomic profiles [18, 20, 23, 24, 33]. Such molecular subgroups have been correlated with various outcomes of the disease, indicating a biological basis behind the clinical cancer heterogeneity. Molecular subtype classification will allow us to predict standard chemotherapy outcomes and select personalized targeted cancer drugs therapy. Increased knowledge of the molecular subtypes of the different cancer types will shape the future of cancer diagnosis, prognosis, and treatment. In [19], using mRNA expression profile analysis, six robust molecular subtypes of colon cancer are identified via distinct biological pathways and reflect new subgroups of prognoses. These gene signatures represent the molecular heterogeneity of colon cancer. Consecutively, this classification leads to recognizing specific and potentially targeted markers for the different subtypes. Another study uses colorectal cancer tumors from three independent cohorts using transcriptomics data to identify colorectal cancer molecular subtypes [4]. Their results show that identified molecular subtypes are highly correlated with microenvironmental signatures. Furthermore, many microarray studies have identified different molecular subtypes of breast cancer. In [17], using the genomic grade (a tumor description focused on how cancer cells and tissues appear under a microscope and how rapidly cancer cells are likely to develop and spread), two clinically distinct estrogen receptor (ER)-positive molecular subtypes are identified. Using gene expression microarrays from patients with triple-negative breast cancer (TNBC), the authors in [31] identified seven molecular subtypes that predict high versus low pathologic complete response (pCR) rates after neoadjuvant chemotherapy. Classification of the subtype will inspire novel personalized medicine approaches for TNBC patients.

Advances in machine learning techniques and image processing, and the availability of large-scale whole-slide histology image databases have led to the development of complex deep learning models in computational pathology [8, 15, 16, 25, 31]. Compared with human inspection, there is tremendous potential for computerized image processing to increase performance, accuracy, and consistency in the diagnosis and prognosis. Computational histological image analysis has been quite promising for diagnosis and biomarker discovery in several cancer types [5, 6, 10, 22, 32, 34–36]. Computational histopathological analysis can identify candidate regions that require more diagnosis evaluation and so can be used in the prognosis of various cancer types [5, 26–28, 34]. Many studies showed that CNNs outperform conventional machine learning approaches for classifying high-resolution images [7, 8, 14]. Also, compared to conventional classification approaches that rely on feature extraction methods, the CNN approaches for classifying histological images reduce the need for extensive domain knowledge from the machine learning expert, as the network learns all relevant features by itself when sufficient data is available [2, 9, 13, 21, 29].

The Cancer Genome Atlas (TCGA) provides a comprehensive set of genomic data for a large population of patients arising from 33 types of cancer and their matched histopathological images for solid tumor samples. Therefore, using these huge data sets, TCGA is able to help studies to find the relationship between morphological and genomic features, which is an important scientific problem. Using the complete set of tumors in TCGA, a comprehensive integrative molecular analysis was conducted to provide a first systematic view of the molecular factors that differentiate the various neoplasms [12]. This PanCancer study included 11,286 tumor samples of 33 types of cancer for which molecular data are available from multiple assay platforms. Using iCluster [30], they identified 28 molecular subtypes and associated them to histological subtypes. This study applied the hierarchical clustering of tumor samples using chromosome arm-level aneuploidy, DNA hypermethylation, microRNA, and Protein to identify 10, 25, 15 and, 8 molecular subtypes, respectively. Also, it uses unsupervised consensus clustering of tumor samples by mRNA expression data to identify 25 molecular subtypes. Then, the clustering of cluster assignment (COCA) algorithm was used to determine the overlap of platform-specific memberships of each of the five molecular platforms to identify the final 28 molecular subtypes.

In this work, we present a new PanCancer classification of tumor histological images using deep learning to predict molecular subtypes of different samples. We used whole-slide hematoxylin and eosin (H&E) images of solid tumors from TCGA that are labeled by the identified molecular subtypes from the integrative analyses introduced in [12]. We employed transfer learning to the Google Inception V3 model by modifying its last layer to classify H&E images and predict molecular subtypes status. Also, we embedded high dimensional RNA-Seq data into 2-D images and developed an integrative deep learning framework to classify TCGA samples according to their molecular subtypes. Extracted features from the H&E images and gene expression profiles are the inputs of this integrative model. Using our integrative model, we provide new insights into the relationship between cancer tissue morphology and gene expression.

2 DATASETS

Our PanCancer study includes 22,484 H&E-stained whole-slide images and the matched normalized-level 3 RNA-Seq gene expression data from 9,383 solid tumor samples that were directly downloaded from the TCGA portal.

For whole-slide images, we used SVS files from primary solid tumors with 20X and 40X magnification and resized the 40X images to 20X. We assigned 70% of the slides to the training set and 30% of them to the test set. We applied the method in [37] to remove the background from the SVS files. The whole-slide images are very large (ranging from 15,000 to 100,000 pixels in width and height) compared to the practically feasible dimensions of the CNN. As a result, we partitioned images into non-overlapping tiles of 512×512 pixels. Overall, we generated 40,994,742 tiles from 22,484 H&E slides for 32 types of cancer. Also, we downloaded RNAseq gene expression data of the 32 cancer types from TCGA PanCancer Atlas. The data was preprocessed by log-transformation of expression values, and genes that have no expression values across all the samples were removed. After preprocessing, we sorted the genes by variance and kept the top 10,000 genes with the highest variances. We then reshaped this vector of size 10,000 into a 100×100 pixel 2D image where the order of genes in the original vector was

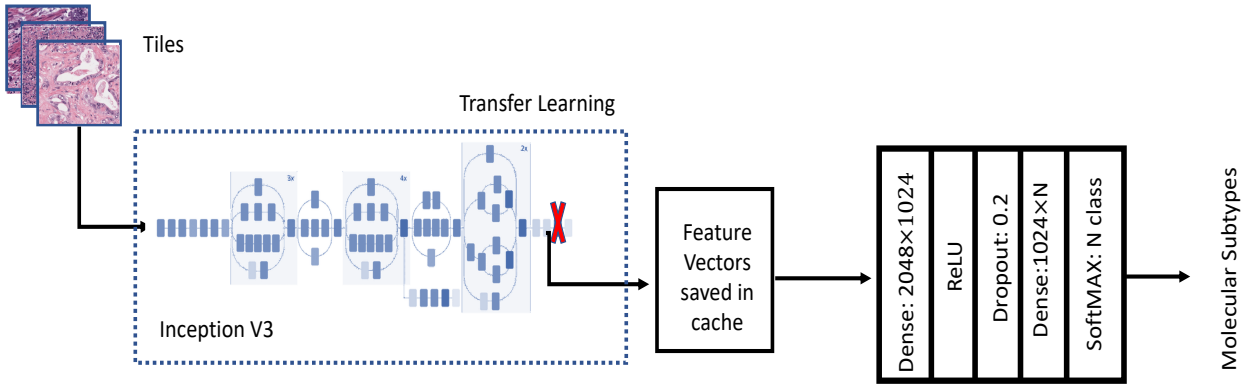

Figure 1: Molecular subtype classification pipeline using transfer learning and histological images. The proposed CNN model classifies histological slides from TCGA by molecular subtypes status. The extracted features from pre-trained Inception V3 are fed into a dense layer with 1024 neurons.

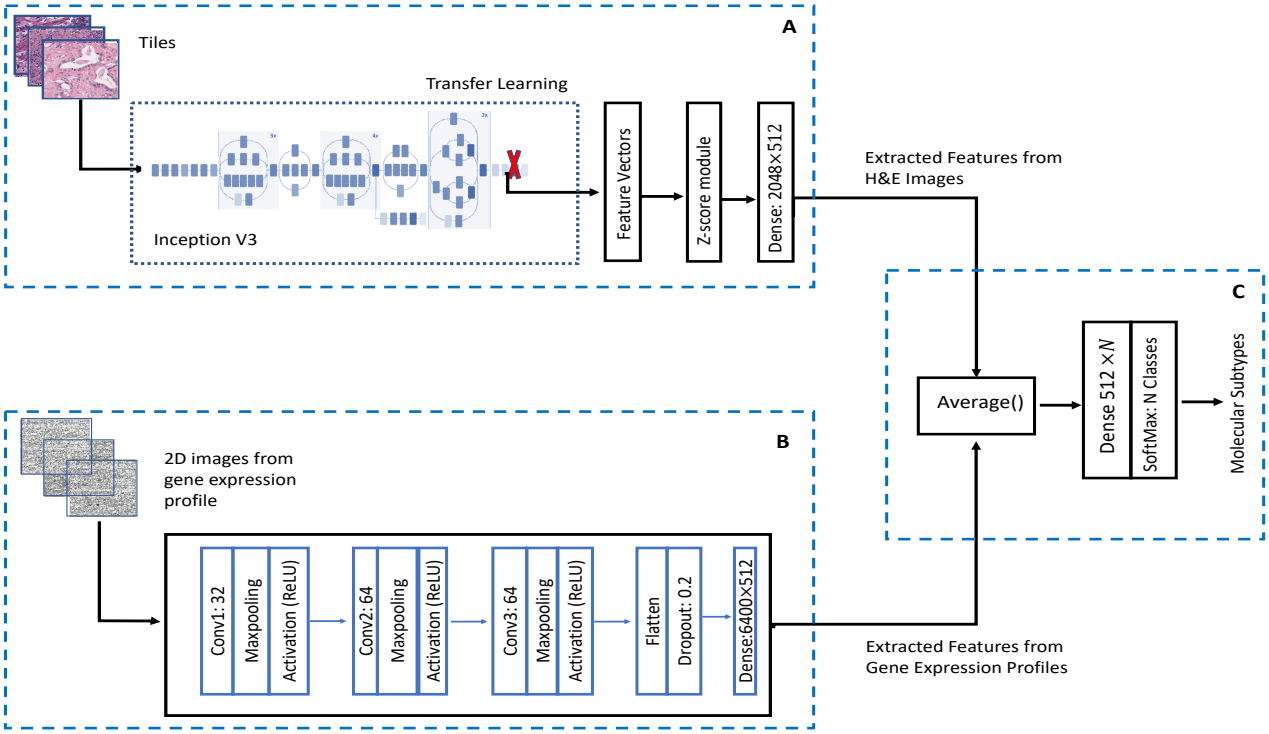

Figure 2: Integrative workflow for molecular subtypes classification using histological images and gene expression profiles. The image features are extracted from histological images using a pre-trained Inception V3, Z-score module, and a dense layer (block A). The genomic features are extracted using a CNN model with three convolutional layers and one dense layer (block B). The extracted features from images and genes are integrated together and trained using a dense layer to predict the molecular subtypes of samples (block C).

preserved in a row-by-row manner. Therefore, each sample has one image of gene expression values. The gene expression images were normalized to make sure the range is between 0 and 255.

We labeled gene expression images and histological slides/tiles using molecular subtypes labels of the samples from the integrative analysis in [12].

3 METHODS

In this section, first, we describe the implementation of the CNN architecture used for analyzing the H&E slides for molecular subtype classification. Then, we introduce an integrative framework that inputs image features from H&E images and gene expression features from RNAseq data to predict patient molecular subtype status.

3.1 CNN model for molecular subtyping using H&E images

In this section, we describe the CNN architecture we developed to analyze the H&E slides from TCGA for molecular subtype classification. Training a complex CNN model for analyzing high-resolution images is a time-consuming task. One way to shorten the training time is to use transfer learning. Transfer learning reuses the weights (parameters) from a pre-trained model on another dataset (e.g., ImageNet for image recognition tasks) and only retains one or few final layers on the new dataset of interest. Also, training a complex CNN model with many layers from scratch requires the availability of many training samples. Well-annotated training datasets are difficult and costly to obtain, especially in the medical imaging domain. Extracting features using a pre-trained network, rather than learning the features from scratch, is a common practice in the medical imaging domain. As a result, transfer learning has been widely used in digital pathology [31].

In this study for molecular subtype classification using H&E slides, we implemented a CNN architecture based on the Google Inception V3 model. We used the exact structure as the Inception V3 model and only deleted the last layer (the classification layer). We used transfer learning and froze all the parameters of the pre-trained Inception V3 model. Fully training is computationally very expensive hence we only relied on transfer learning [22]. As shown in Figure 1, the generated 512×512 pixel tiles are inputs to the Inception V3 network in a forward pass, and we stored the values of the last fully connected layer as vectors of 2,048 floating-point values. We shuffled the stored vectors from a similar holdout group and assigned them to TFRecords in groups of 10,000. Theses TFRecords were used as inputs to a fully connected layer with 1024 neurons, and its output was encoded to a one-hot-encoded vector. We ran 100,000 iterations for each training simulation in batches of 512 samples, with a 20% dropout. We used the cross-entropy loss function, and Adam optimizer with the learning rate of 0.01 to update the parameters. A SoftMax function was used to derive class probabilities.

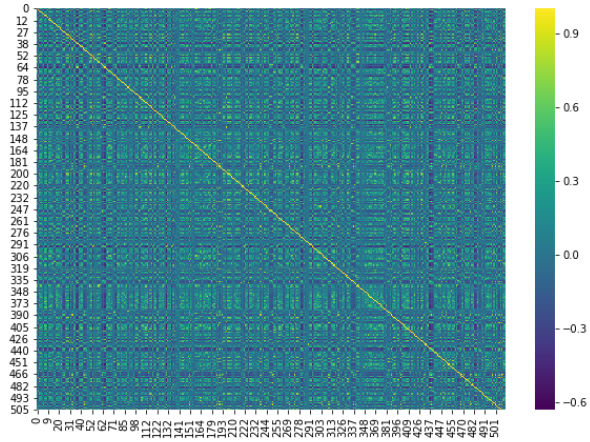

Figure 3: The heatmap of extracted features from H&E images and gene expression data. Two feature vectors are highly correlated.

3.2 Integrative model for molecular subtype classification using H&E images and expression profiles

We developed an integrative model that combines the extracted features from the H&E images using the Inception V3 network and gene expression features from another CNN model to identify molecular subtype status of solid tumors (Figure 1). Our integrative model is divided into three blocks: H&E images feature extraction (block A), gene expression data feature extraction (block B), and integrative classifier network (block C).

As shown in Figure 2 (block A), the output of the last fully connected layer of Inception V3 (with 2,048 neurons) is fed to the Z-score module. The hypothesis behind Z-score is to describe any data points by finding out how they relate to the standard deviation and mean of the data points' group. By calculating the Z-score ($Z = (x - \mu)/\sigma$, where x is an extracted feature vector of a tile, and μ and σ are the mean and standard deviation of features across the corresponding slide), we looked for tiles where their extracted features are close to the mean of the distribution of all tiles' features of a slide. We kept only three tiles from each slide that have the smallest absolute Z-score for their extracted features. Then we calculated the mean of the features of the three tiles to have one vector of 2,048 floating-point values as a feature vector for each sample. The samples' feature vectors of size 2048 are the inputs of another fully connected layer that has 512 neurons in its output. Therefore, at the end for each sample, we have a vector of size 512 as its extracted H&E image features.

We also extracted features from gene expression data using a CNN network, as shown in Figure 2, block B. The CNN consists of three convolutional layers and one dense layer. We used the cross-entropy loss function, and Adam optimizer with the learning rate of 0.001 to update the parameters. The first convolutional layer has 32

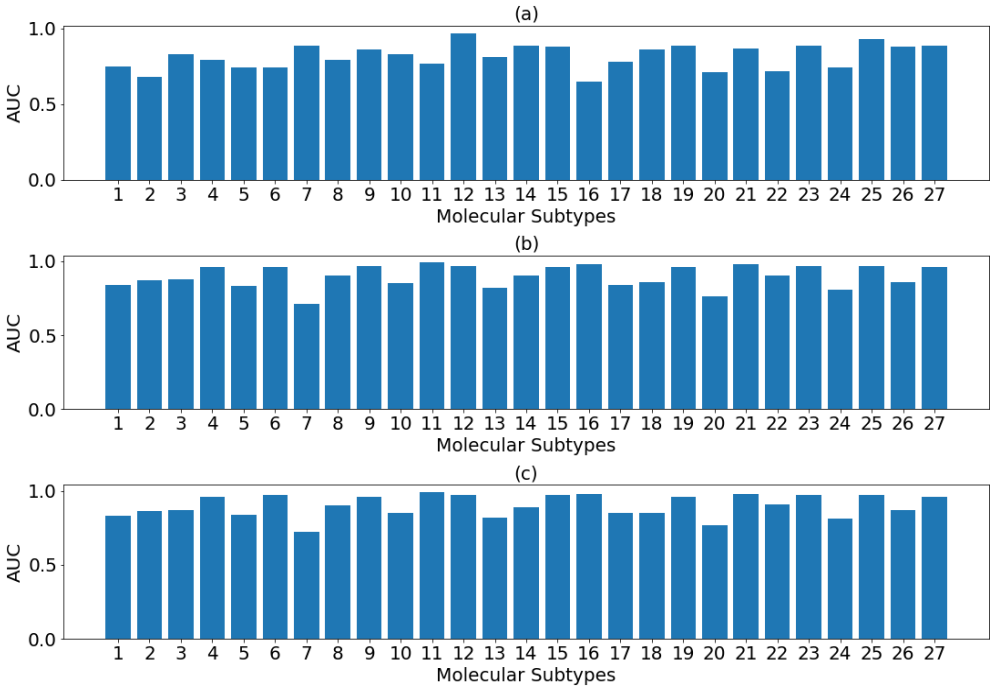

Figure 4: AUCs for a) per-tile analysis, b) per-slide analysis c) per-patient analysis, using histological images for all predicted classes.

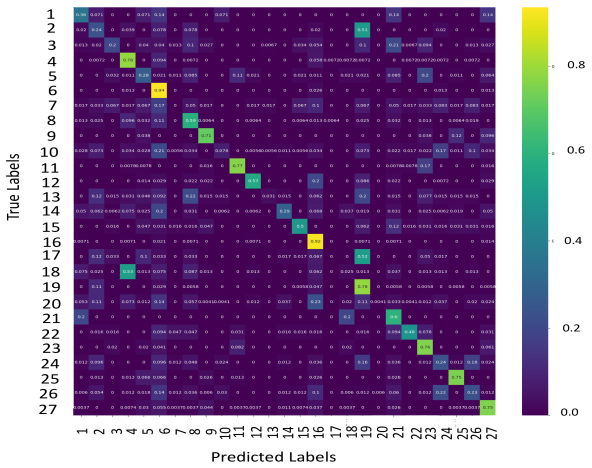

Figure 5: Confusion Matrix from per-patient analysis from the SoftMax of the image classification model.

filters, while the second and third layers have 64 filters, respectively. The size of the filters is 3×3. Each convolutional layer is followed by a max-pooling layer and a ReLU activation layer. Following the convolutional layers is a flatten layer and a dropout layer with a dropout rate of 0.2. We then added one dense layer that its output layer has 512 neurons. Therefore, for each sample, we have a vector of 512 floating-point values as its genomic features.

Then as can be seen in Figure 2, block C, we calculated the average between image features and genomic features of each sample. Doing this for each sample, we obtained a vector of 512 floating-point values as its final integrative extracted features. We then added one dense layer with a size of 512. For this dense layer, we used the cross-entropy loss function, and Adam optimizer with the learning rate of 0.01 to update the parameters. The output layer has 27 neurons, which are activated by the SoftMax function.

3.3 Relationship between H&E image features and gene expression features

Using a correlation coefficients matrix between H&E image features from the Inception V3 model and gene expression features from the CNN, we observed that they are correlated. The heatmap of Pearson product-moment correlation coefficients, R_{ij} is shown in Figure 3 ($R_{ij} = C_{ij}/\sqrt{C_{ii} \times C_{jj}}$, where C is the covariance between the image features from Inception V3 and the gene expression features from the CNN). The high correlations between H&E image features and gene expression features suggest that, as well as genomic data, H&E images can also be used for molecular subtyping. We will show that despite this high correlation, integrating H&E images and genomic data can improve molecular subtype classification indicating that both images and genes can contribute to identify molecular subtypes.

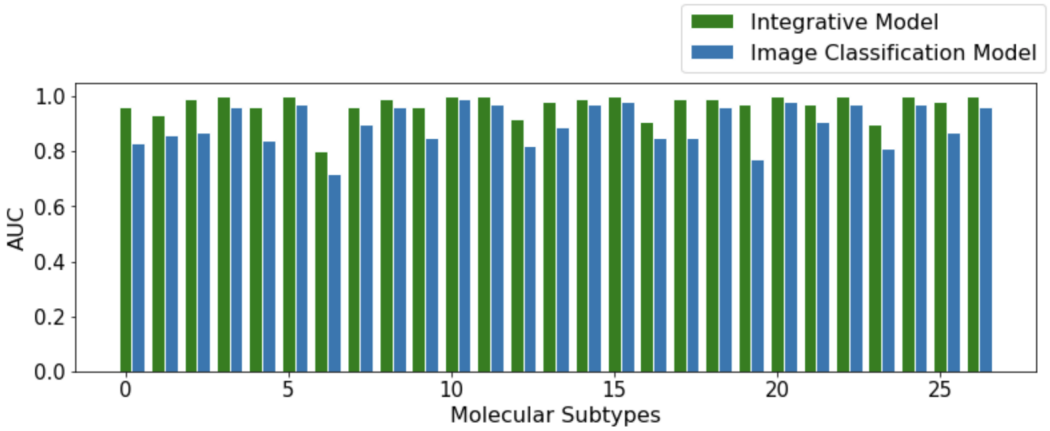

Figure 6: AUCs for molecular subtypes classification using the integrative model and the image classification model.

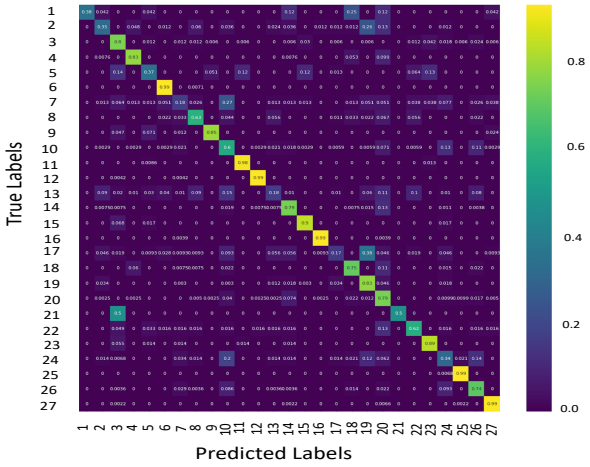

Figure 7: Confusion Matrix from the SoftMax of the integrative model.

3.4 Programming details

All codes are written with Python, and the neural network architectures are implemented with TensorFlow. All transfer learning analysis was performed on the Google Cloud Platform (GCP). We used Kubernetes, Datastore, Cloud Storage, and Pub/sub on GCP and up to 1000 compute instances (each 16 vCPUs and 60GB memory) and up to 4000 Kubernetes pods.

4 RESULTS

4.1 Classification using H&E images

As we discussed in the previous section, we applied our CNN algorithm to classify tumor slides according to their molecular subtypes using H&E slide images. We calculated the per-tile classification receiver operating characteristics (ROCs) based on thresholding SoftMax probabilities (Figure 4 (a)). To obtain the molecular subtype prediction per slide, for each slide, we calculated means of the tile-level SoftMax probabilities for all classes. We assigned a class to each slide based on the maximum value of the means of the SoftMax probabilities. Using these per-slide predictions, we calculated ROCs for per-slide classification (Figure 4 (b)). Since each patient has multiple slides, we followed the same approach in per slide prediction for deriving the label of patients from slide-level analysis (figure 4 (c)). Calculating the area under the curve (AUC) of ROCs (one versus all other classes), we can see that the proposed CNN model accurately classifies the tiles for most molecular subtypes (AUC micro-average = 0.83 ; macro-average = 0.82) (Figure 4 (a)). We also calculated AUC for per-slide and per-patient analysis. As you can see in Figures 4 (b) and 4 (c), the CNN model yielded a very high accuracy for slide-level (AUC micro-average = 0.90 ; macro-average = 0.90), and patient-level classification (AUC micro-average = 0.90 ; macro-average = 0.90). The improved classification accuracy from the aggregations at the slide level or patient-level compared to the tile-level is expected, since our subtype annotations are from bulk measurements, and we do not have annotations at the tile level. Some tiles can belong to not cancer cell areas, resulting in misclassification at the tile level. As you can see in Figure 4, in most cases, the proposed algorithm can correctly classify the molecular subtypes.

The confusion matrix shows the relationships between different classes and how errors are distributed. From patient-level confusion matrix (Figure 5), we can see most classes are classified correctly. The confusion matrix indicates that those classes with lower AUC are falsely classified to another class (Figures 4 and 5).

The prediction accuracies are variable across molecular subtypes. This is partially due to sample sizes as the molecular subtypes with higher AUCs have more histological images. Also, tumors originating from the same organ tend to belong to the same molecular subtype. Whereas classes with lower AUC values tend to be heterogeneous in cancer types or tissue of origin.

4.2 Classification using H&E images and gene expression

We calculated AUCs for the classification of solid tumor samples according to their molecular subtypes using our developed integrative model. As we discussed before, these molecular subtypes arising from the 33 different tumor types analyzed across at least four different TCGA platforms that assayed tumor DNA (DNA methylation and copy number), and RNA (mRNA and microRNA sequencing) [35]. As we expected, adding genomic data and integrating them with histological images leads to higher AUCs for most of the classes (AUC micro-average= 0.99; macro-average=0.97) (Figure 6). Figures 6 and 7 show that our integrative approach yielded more accurate classification results for all classes. In Figure 6, we see that for the integrative model, all classes have a higher AUC compared to the H&E image classification model. In Figure 7, the confusion matrix shows that most classes are classified correctly when we use our developed integrative model.

5 CONCLUSION

Pathologists analysis and interpretation of tissue sections stained with H&E is an important component of disease evaluation. The phenotypic information found in histology slides can be used for prognosis as well as providing diagnostic information. Automated computational histopathological imaging systems have shown great promise in the detection and identification of new cancer biomarkers for these large data. The CNN-based deep learning approaches significantly outperform other conventional approaches in a variety of tasks in digital pathology. Also, TCGA project offers a comprehensive set of genomics and clinical outcome data for large populations of patients with more than 30 types of cancer and a large collection of related histopathological images for solid tumor samples. Using these comprehensive datasets and a transfer learning technique, we implemented a PanCancer molecular subtype classification of TCGA H&E slides.

We applied the Inception V3 CNN to histopathological images of 32 cancer types and predicted their molecular subtype class as defined in [12], and were able to classify images with very high accuracies (micro-average= 0.90; macro-average=0.90). Our results show that imaging data is highly predictive of molecular subtypes and suggests that these two data modalities are highly correlated. To directly assess the associations between imaging and molecular subtypes, we independently trained a second neural network only on gene expression data and used it as a feature extractor for comparison to image features derived from the final layer of the Inception model. We found that the genomic features are highly correlated with image features, suggesting that these two data modality are highly interchangeable for representing molecular subtypes.

To determine if the extracted features combined are more predictive of molecular subtypes, we constructed an integrative model by averaging genomic and image features and classifying them using a single dense layer. Our model was able to classify samples with a very high accuracy (micro-average= 0.99; macro-average=0.97). This result suggests that integration of the two data modalities after an initial embedding provides higher classification accuracies.

In this PanCancer study histology images and gene expression of 9,383 solid tumor samples from 32 cancer types were analyzed and embedded into a low-dimentional space to predict their corresponding molecular subtypes. Further studies to determine the associations between these features and clinical data may provide new opportunities for extracting PanCancer biomarkers independent of cancer type.

REFERENCES

[1] Kimberly H. Allison and George W. Sledge. 2014. Heterogeneity and cancer. *Oncology (Williston Park, N.Y.)* 28, 9 (Sept. 2014), 772–778.

[2] Teresa Araújo, Guilherme Aresta, Eduardo Castro, José Rouco, Paulo Aguiar, Catarina Eloy, António Polónia, and Aurélio Campilho. 2017. Classification of breast cancer histology images using Convolutional Neural Networks. *PLoS ONE* 12, 6 (June 2017), e0177544. https://doi.org/10.1371/journal.pone.0177544

[3] UlyssesJ Balis, Jason Hipp, Thomas Flotte, James Monaco, Jerome Cheng, Anant Madabhushi, Yukako Yagi, Jaime Rodriguez-Canales, Michael Emmert-Buck, MichaelC Dugan, Stephen Hewitt, Mehmet Toner, RonaldG Tompkins, David Lucas, and JohnR Gilbertson. 2011. Computer aided diagnostic tools aim to empower rather than replace pathologists: Lessons learned from computational chess. *J Pathol Inform* 2, 1 (2011), 25. https://doi.org/10.4103/2153-3539.82050

[4] Etienne Becht, Aurélien de Reyniès, Nicolas A. Giraldo, Camilla Pilati, Bénédicte Buttard, Laetitia Lacroix, Janick Selves, Catherine Sautès-Fridman, Pierre Laurent-Puig, and Wolf Herman Fridman. 2016. Immune and Stromal Classification of Colorectal Cancer Is Associated with Molecular Subtypes and Relevant for Precision Immunotherapy. *Clin Cancer Res* 22, 16 (Aug. 2016), 4057–4066. https://doi.org/10.1158/1078-0432.CCR-15-2879

[5] A. H. Beck, A. R. Sangoi, S. Leung, R. J. Marinelli, T. O. Nielsen, M. J. van de Vijver, R. B. West, M. van de Rijn, and D. Koller. 2011. Systematic Analysis of Breast Cancer Morphology Uncovers Stromal Features Associated with Survival. *Science Translational Medicine* 3, 108 (Nov. 2011), 108ra113–108ra113. https://doi.org/10.1126/scitranslmed.3002564

[6] Jun Cheng, Jie Zhang, Yatong Han, Xusheng Wang, Xiufen Ye, Yuebo Meng, Anil Parwani, Zhi Han, Qianjin Feng, and Kun Huang. 2017. Integrative Analysis of Histopathological Images and Genomic Data Predicts Clear Cell Renal Cell Carcinoma Prognosis. *Cancer Res* 77, 21 (Nov. 2017), e91–e100. https://doi.org/10.1158/0008-5472.CAN-17-0313

[7] Dan Cireşan, Ueli Meier, and Juergen Schmidhuber. 2012. Multi-column Deep Neural Networks for Image Classification. *arXiv:1202.2745 [cs]* (Feb. 2012). http://arxiv.org/abs/1202.2745 arXiv: 1202.2745.

[8] Dan C. Cireşan, Alessandro Giusti, Luca M. Gambardella, and Jürgen Schmidhuber. 2013. Mitosis detection in breast cancer histology images with deep neural networks. *Med Image Comput Comput Assist Interv* 16, Pt 2 (2013), 411–418. https://doi.org/10.1007/978-3-642-40763-5_51

[9] Okyaz Eminaga, Mahmood Abbas, Yuri Tolkach, Rosalie Nolley, Christian Kunder, Axel Semjonow, Martin Boegemann, Andreas Loening, James Brook, and Daniel Rubin. 2019. Biologic and Prognostic Feature Scores from Whole-Slide Histology Images Using Deep Learning. *arXiv:1910.09100 [cs, q-bio]* (Oct. 2019). http://arxiv.org/abs/1910.09100 arXiv: 1910.09100.

[10] Tunc Gultekin, Can Fahrettin Koyuncu, Cenk Sokmensuer, and Cigdem Gunduz-Demir. 2015. Two-Tier Tissue Decomposition for Histopathological Image Representation and Classification. *IEEE Trans. Med. Imaging* 34, 1 (Jan. 2015), 275–283. https://doi.org/10.1109/TMI.2014.2354373

[11] M.N. Gurcan, L.E. Boucheron, A. Can, A. Madabhushi, N.M. Rajpoot, and B. Yener. 2009. Histopathological Image Analysis: A Review. *IEEE Rev. Biomed. Eng.* 2 (2009), 147–171. https://doi.org/10.1109/RBME.2009.2034865

[12] Katherine A Hoadley, Christina Yau, Toshinori Hinoue, Denise M Wolf, Alexander J Lazar, Esther Drill, Ronglai Shen, Alison M Taylor, Andrew D Cherniack, Vésteinn Thorsson, et al. 2018. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* 173, 2 (2018), 291–304.

[13] Vijaya B. Kolachalama, Priyamvada Singh, Christopher Q. Lin, Dan Mun, Mostafa E. Belghasem, Joel M. Henderson, Jean M. Francis, David J. Salant, and Vipul C. Chitalia. 2018. Association of Pathological Fibrosis With Renal Survival Using Deep Neural Networks. *Kidney International Reports* 3, 2 (March 2018), 464–475. https://doi.org/10.1016/j.ekir.2017.11.002

[14] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 1097–1105. http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[15] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2017. ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60, 6 (May 2017), 84–90. https://doi.org/10.1145/3065386

[16] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A. W. M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. 2017. A Survey on Deep Learning in

Medical Image Analysis. *Medical Image Analysis* 42 (Dec. 2017), 60–88. https://doi.org/10.1016/j.media.2017.07.005 arXiv: 1702.05747.

[17] Sherene Loi, Benjamin Haibe-Kains, Christine Desmedt, Françoise Lallemand, Andrew M. Tutt, Cheryl Gillet, Paul Ellis, Adrian Harris, Jonas Bergh, John A. Foekens, Jan G.M. Klijn, Denis Larsimont, Marc Buyse, Gianluca Bontempi, Mauro Delorenzi, Martine J. Piccart, and Christos Sotiriou. 2007. Definition of Clinically Distinct Molecular Subtypes in Estrogen Receptor–Positive Breast Carcinomas Through Genomic Grade. *JCO* 25, 10 (April 2007), 1239–1246. https://doi.org/10.1200/JCO.2006.07.1522

[18] Gautam K. Malhotra, Xiangshan Zhao, Hamid Band, and Vimla Band. 2010. Histological, molecular and functional subtypes of breast cancers. *Cancer Biology & Therapy* 10, 10 (Nov. 2010), 955–960. https://doi.org/10.4161/cbt.10.10.13879

[19] Laetitia Marisa, Aurélien de Reyniès, Alex Duval, Janick Selves, Marie Pierre Gaub, Laure Vescovo, Marie-Christine Etienne-Grimaldi, Renaud Schiappa, Dominique Guenot, Mira Ayadi, Sylvain Kirzin, Maurice Chazal, Jean-François Fléjou, Daniel Benchimol, Anne Berger, Arnaud Lagarde, Erwan Pencreach, Françoise Piard, Dominique Elias, Yann Parc, Sylviane Olschwang, Gérard Milano, Pierre Laurent-Puig, and Valérie Boige. 2013. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med.* 10, 5 (2013), e1001453. https://doi.org/10.1371/journal.pmed.1001453

[20] H. Masuda, K. A. Baggerly, Y. Wang, Y. Zhang, A. M. Gonzalez-Angulo, F. Meric-Bernstam, V. Valero, B. D. Lehmann, J. A. Pietenpol, G. N. Hortobagyi, W. F. Symmans, and N. T. Ueno. 2013. Differential Response to Neoadjuvant Chemotherapy Among 7 Triple-Negative Breast Cancer Molecular Subtypes. *Clinical Cancer Research* 19, 19 (Oct. 2013), 5533–5540. https://doi.org/10.1158/1078-0432.CCR-13-0799

[21] Pooya Mobadersany, Safoora Yousefi, Mohamed Amgad, David A. Gutman, Jill S. Barnholtz-Sloan, José E. Velázquez Vega, Daniel J. Brat, and Lee A. D. Cooper. 2018. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci USA* 115, 13 (March 2018), E2970–E2979. https://doi.org/10.1073/pnas.1717139115

[22] Javad Noorbakhsh, Saman Farahmand, Ali Foroughi pour, Sandeep Namburi, Dennis Caruana, David Rimm, Mohammad Soltanieh-ha, Kourosh Zarringhalam, and Jeffrey H. Chuang. 2020. *Deep learning-based cross-classifications reveal conserved spatial behaviors within tumor histological images.* preprint. Bioinformatics. https://doi.org/10.1101/715656

[23] Seho Park, Ja Seung Koo, Min Suk Kim, Hyung Seok Park, Jun Sang Lee, Jong Seok Lee, Seung Il Kim, and Byeong-Woo Park. 2012. Characteristics and outcomes according to molecular subtypes of breast cancer as classified by a panel of four biomarkers using immunohistochemistry. *The Breast* 21, 1 (Feb. 2012), 50–57. https://doi.org/10.1016/j.breast.2011.07.008

[24] Amanda I. Phipps, Paul J. Limburg, John A. Baron, Andrea N. Burnett-Hartman, Daniel J. Weisenberger, Peter W. Laird, Frank A. Sinicrope, Christophe Rosty, Daniel D. Buchanan, John D. Potter, and Polly A. Newcomb. 2015. Association Between Molecular Subtypes of Colorectal Cancer and Patient Survival. *Gastroenterology* 148, 1 (Jan. 2015), 77–87.e2. https://doi.org/10.1053/j.gastro.2014.09.038

[25] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. 2015. ImageNet Large Scale Visual Recognition Challenge. *arXiv:1409.0575 [cs]* (Jan. 2015). http://arxiv.org/abs/1409.0575 arXiv: 1409.0575.

[26] Edmond Sabo, Andrew H Beck, Elizabeth A Montgomery, Baishali Bhattacharya, Patricia Meitner, Ji Yi Wang, and Murray B Resnick. 2006. Computerized morphometry as an aid in determining the grade of dysplasia and progression to adenocarcinoma in Barrett's esophagus. *Lab Invest* 86, 12 (Dec. 2006), 1261–1271. https://doi.org/10.1038/labinvest.3700481

[27] Olcay Sertel, Jun Kong, Umit V. Catalyurek, Gerard Lozanski, Joel H. Saltz, and Metin N. Gurcan. 2009. Histopathological Image Analysis Using Model-Based Intermediate Representations and Color Texture: Follicular Lymphoma Grading. *J Sign Process Syst Sign Image Video Technol* 55, 1-3 (April 2009), 169–183. https://doi.org/10.1007/s11265-008-0201-y

[28] O. Sertel, J. Kong, H. Shimada, U.V. Catalyurek, J.H. Saltz, and M.N. Gurcan. 2009. Computer-aided prognosis of neuroblastoma on whole-slide images: Classification of stromal development. *Pattern Recognition* 42, 6 (June 2009), 1093–1103. https://doi.org/10.1016/j.patcog.2008.08.027

[29] Mary Shapcott, Katherine J. Hewitt, and Nasir Rajpoot. 2019. Deep Learning With Sampling in Colon Cancer Histology. *Front. Bioeng. Biotechnol.* 7 (March 2019), 52. https://doi.org/10.3389/fbioe.2019.00052

[30] Ronglai Shen, Qianxing Mo, Nikolaus Schultz, Venkatraman E. Seshan, Adam B. Olshen, Jason Huse, Marc Ladanyi, and Chris Sander. 2012. Integrative Subtype Discovery in Glioblastoma Using iCluster. *PLoS ONE* 7, 4 (April 2012), e35236. https://doi.org/10.1371/journal.pone.0035236

[31] Chetan L. Srinidhi, Ozan Ciga, and Anne L. Martel. 2019. Deep neural network models for computational histopathology: A survey. *arXiv:1912.12378 [cs, eess]* (Dec. 2019). http://arxiv.org/abs/1912.12378 arXiv: 1912.12378.

[32] Chao Wang, Thierry Pécot, Debra L Zynger, Raghu Machiraju, Charles L Shapiro, and Kun Huang. 2013. Identifying survival associated morphological features of triple negative breast cancer using multiple datasets. *J Am Med Inform Assoc* 20, 4 (July 2013), 680–687. https://doi.org/10.1136/amiajnl-2012-001538

[33] Lisandra West, Smruti J. Vidwans, Nicholas P. Campbell, Jeff Shrager, George R. Simon, Raphael Bueno, Phillip A. Dennis, Gregory A. Otterson, and Ravi Salgia. 2012. A novel classification of lung cancer into molecular subtypes. *PLoS ONE* 7, 2 (2012), e31906. https://doi.org/10.1371/journal.pone.0031906

[34] Kun-Hsing Yu, Ce Zhang, Gerald J. Berry, Russ B. Altman, Christopher Ré, Daniel L. Rubin, and Michael Snyder. 2016. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun* 7, 1 (Nov. 2016), 12474. https://doi.org/10.1038/ncomms12474

[35] Y. Yuan, H. Failmezger, O. M. Rueda, H. R. Ali, S. Graf, S.-F. Chin, R. F. Schwarz, C. Curtis, M. J. Dunning, H. Bardwell, N. Johnson, S. Doyle, G. Turashvili, E. Provenzano, S. Aparicio, C. Caldas, and F. Markowetz. 2012. Quantitative Image Analysis of Cellular Heterogeneity in Breast Tumors Complements Genomic Profiling. *Science Translational Medicine* 4, 157 (Oct. 2012), 157ra143–157ra143. https://doi.org/10.1126/scitranslmed.3004330

[36] Xiaofan Zhang, Fuyong Xing, Hai Su, Lin Yang, and Shaoting Zhang. 2015. High-throughput histopathological image analysis via robust cell segmentation and hashing. *Medical Image Analysis* 26, 1 (Dec. 2015), 306–315. https://doi.org/10.1016/j.media.2015.10.005