BAYESIAN RANDOM SEMANTIC DATA AUGMENTATION FOR MEDICAL IMAGE CLASSIFICATION

PAPER BY YAOYAO ZHU, XIUDING CAI, XUEYAO WANG AND YU YAO

PRESENTATION BY DIONYSIOS RIGATOS

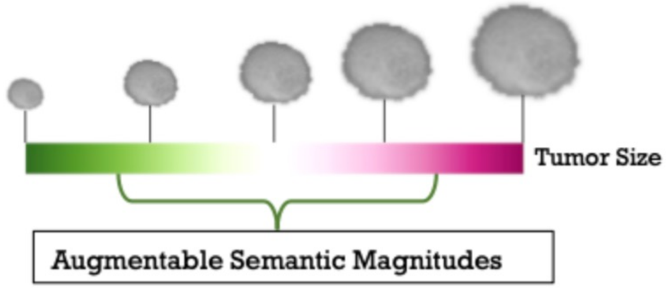
ARXIV URL

MOTIVATION

- Data Augmentation (DA) is a critical regularization technique in Biomedical Image CV tasks
  - Reduces overfitting
  - Increases dataset diversity
  - Medical image have different modalities (i.e. CT scans, MRI, X-RAY)
- Drawbacks with current DA approaches:
  - Image Transformation-Based methods are too simple
  - Automatic methods are too simple
  - Generative methods but are computationally expensive
- Proposed solution:
  - An efficient, plug-and-play semantic augmentation method.

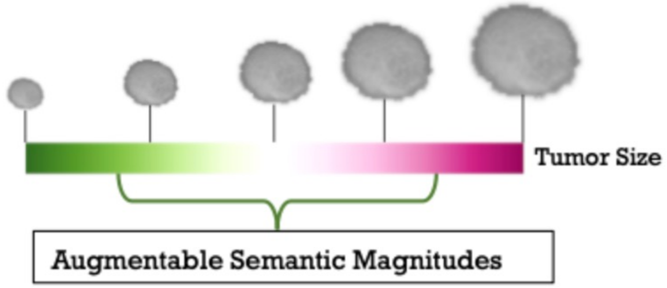
HOW? – SEMANTIC AUGMENTATION

- Addition of semantic magnitude to the original feature…
  - Treated as a random variable, distribution estimated with a variational Bayesian.
  - Randomly sampled from the distribution.

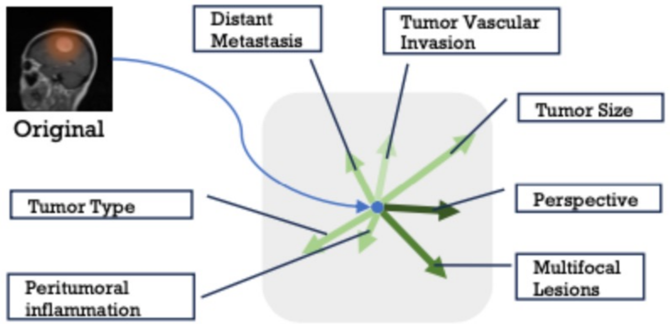


HOW? – SEMANTIC AUGMENTATION

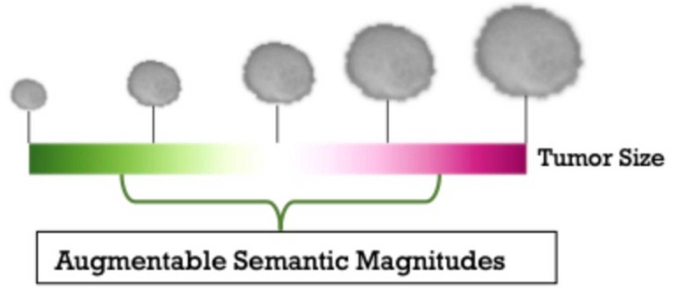
- Addition of semantic magnitude to the original feature…
  - Treated as a random variable, distribution estimated with a variational Bayesian.
  - Randomly sampled from the distribution.

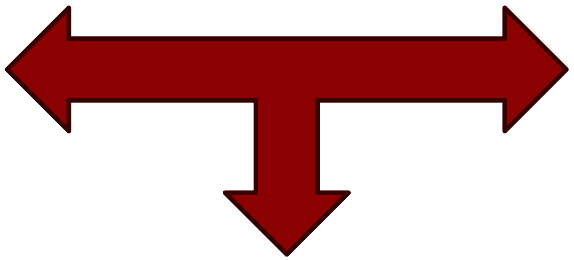


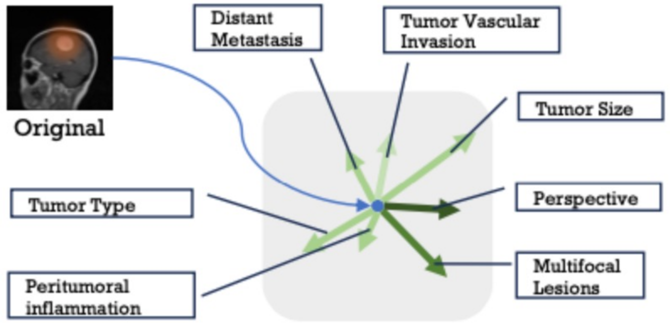
- … in a randomly selected semantic direction.
  - Shifting features along these directions yields new sample features with:
    - Identical class identities.
    - Different semantic context.

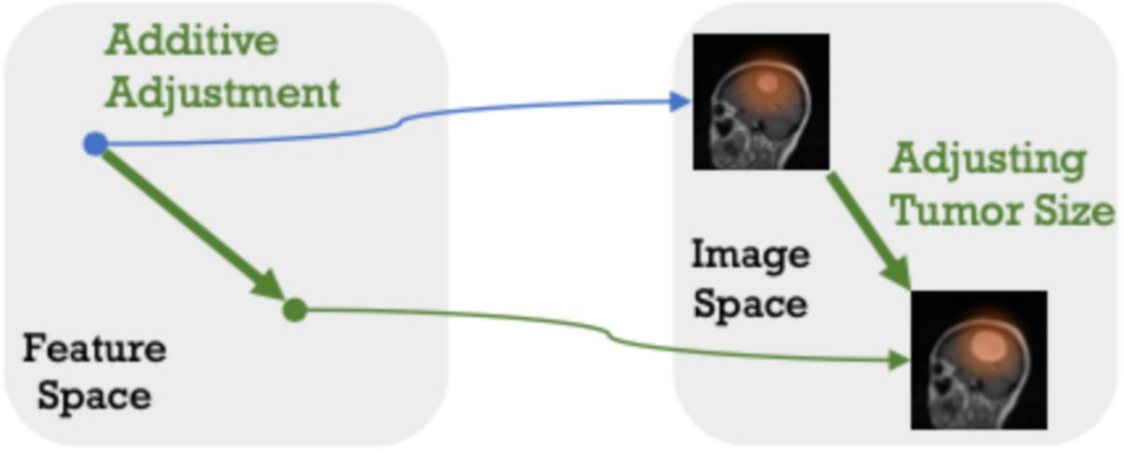


HOW? – SEMANTIC AUGMENTATION



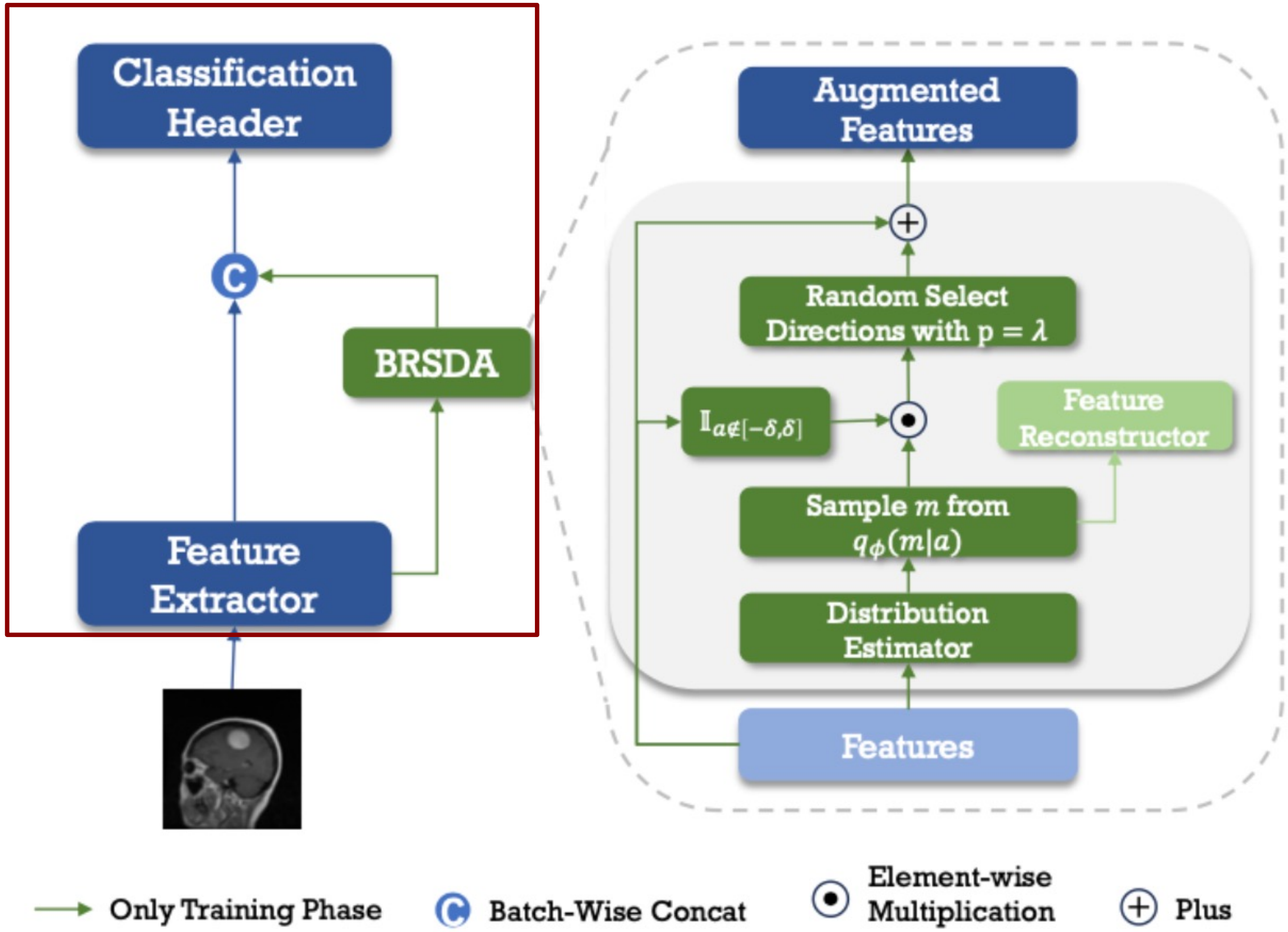






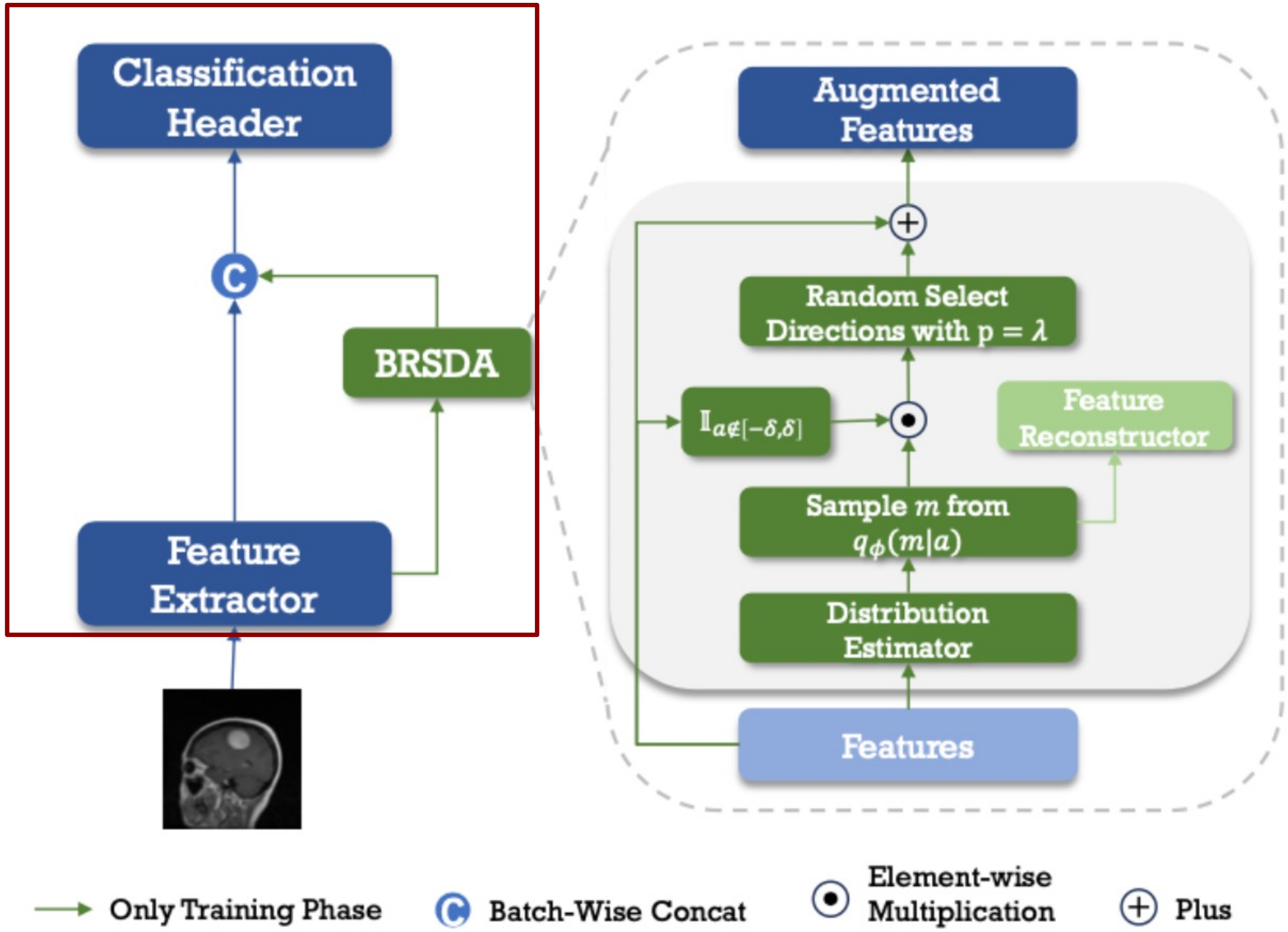
APPROACH – SETUP

$$G_\Theta = f_1 \circ f_2$$



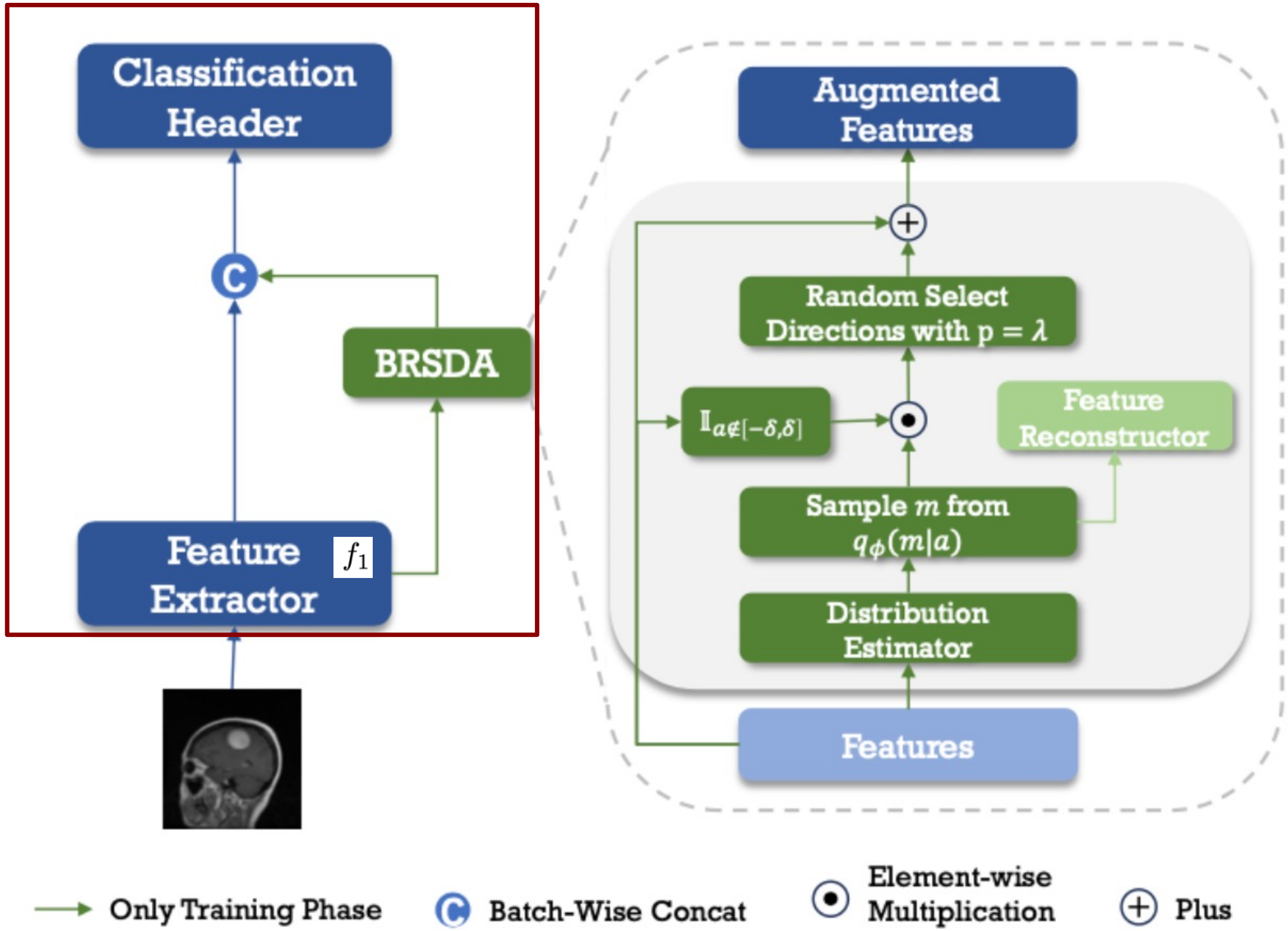
APPROACH – SETUP

$$G_\Theta = f_1 \circ f_2$$



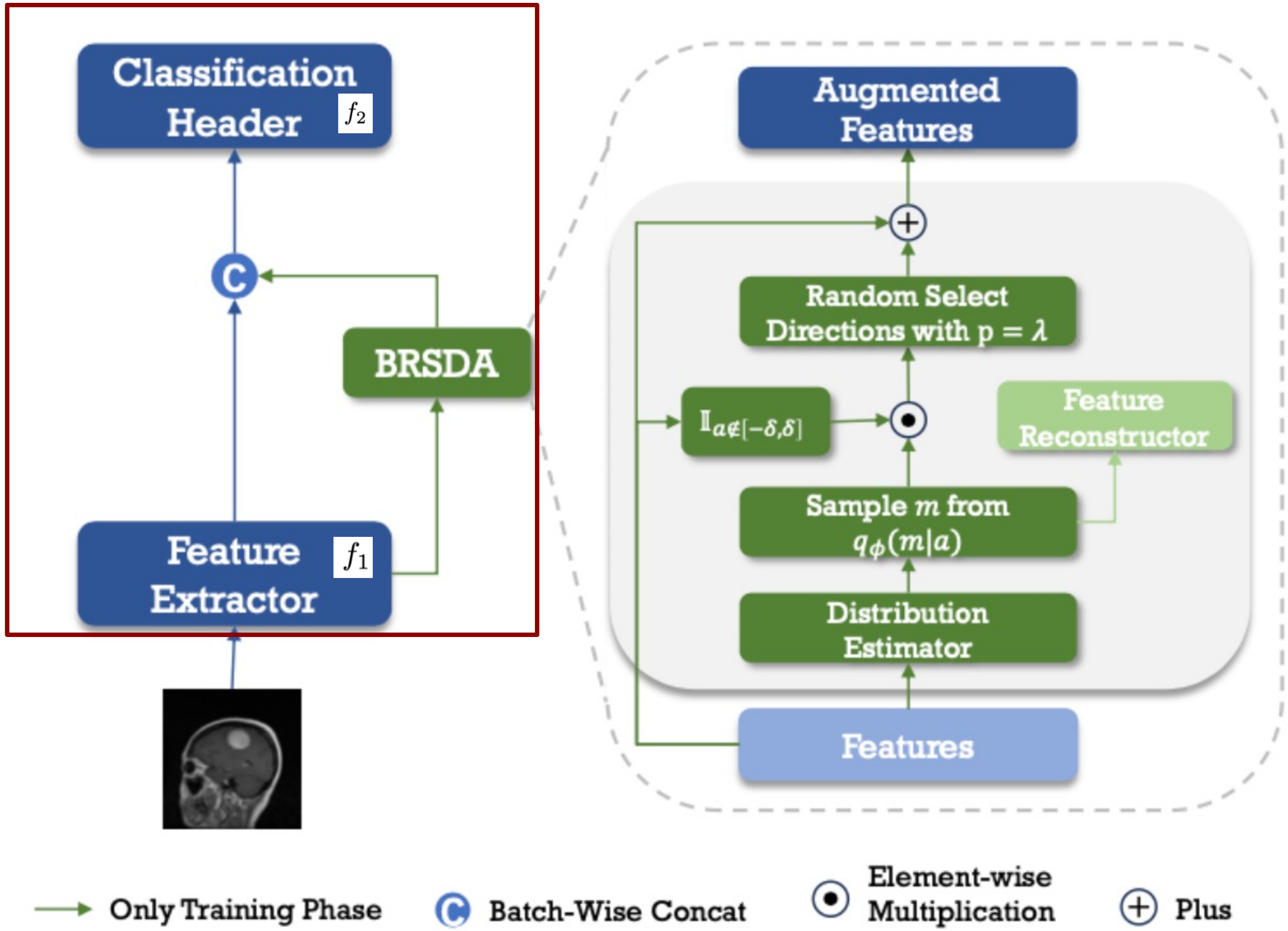
APPROACH – SETUP

$$G_\Theta = f_1 \circ f_2$$



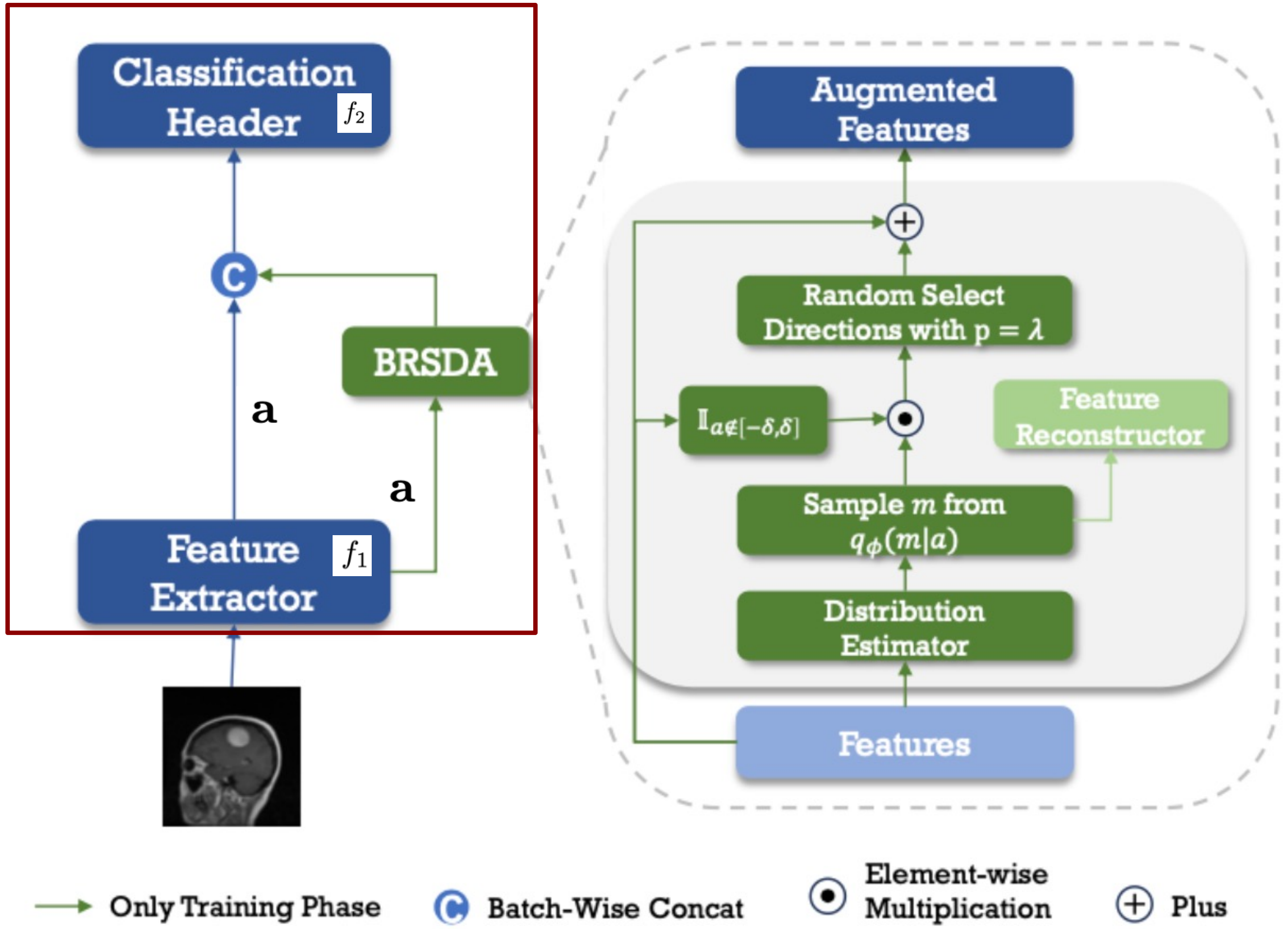
APPROACH – SETUP

$$G_\Theta = f_1 \circ f_2$$



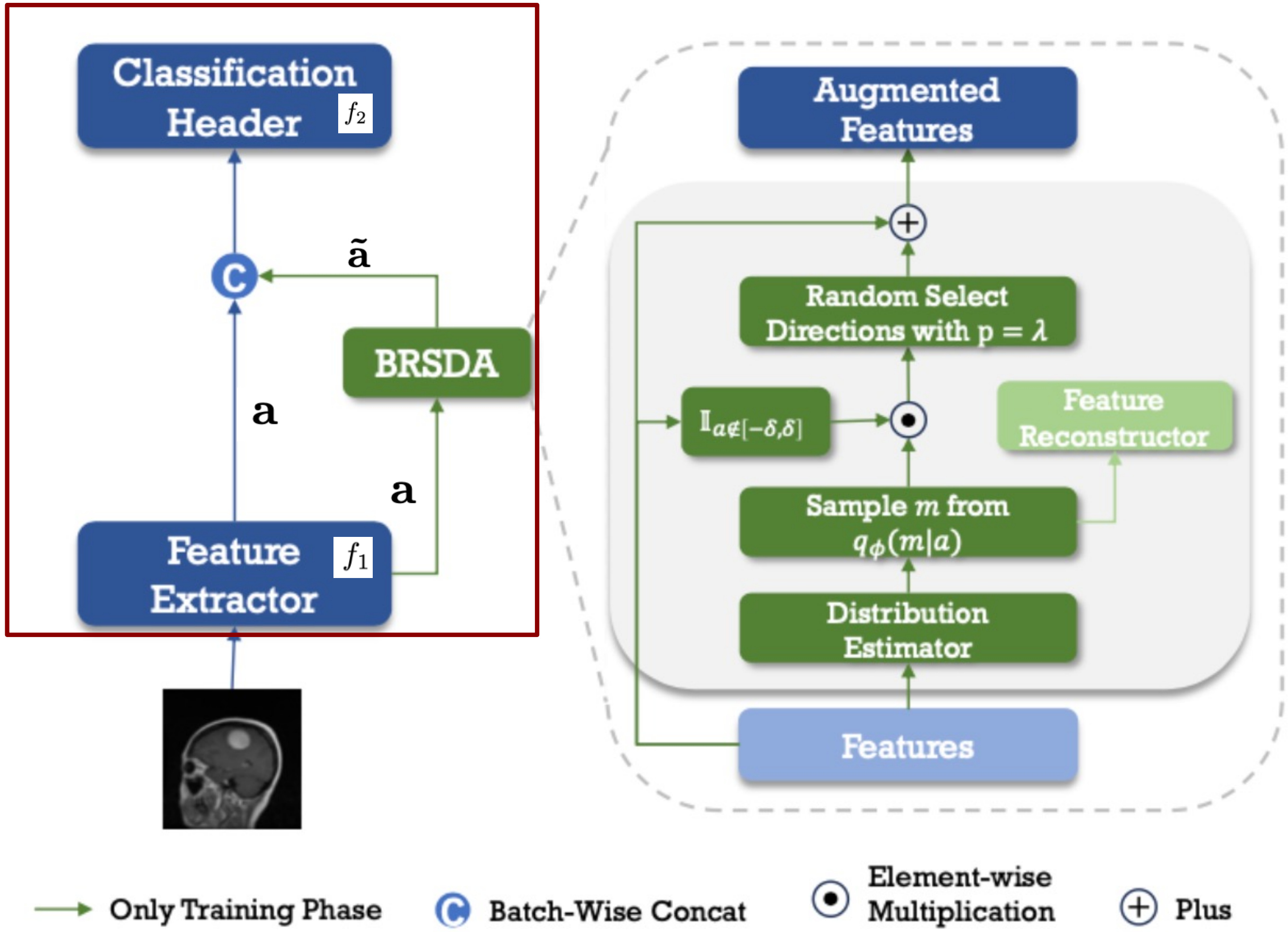
APPROACH – SETUP

$$G_\Theta = f_1 \circ f_2$$

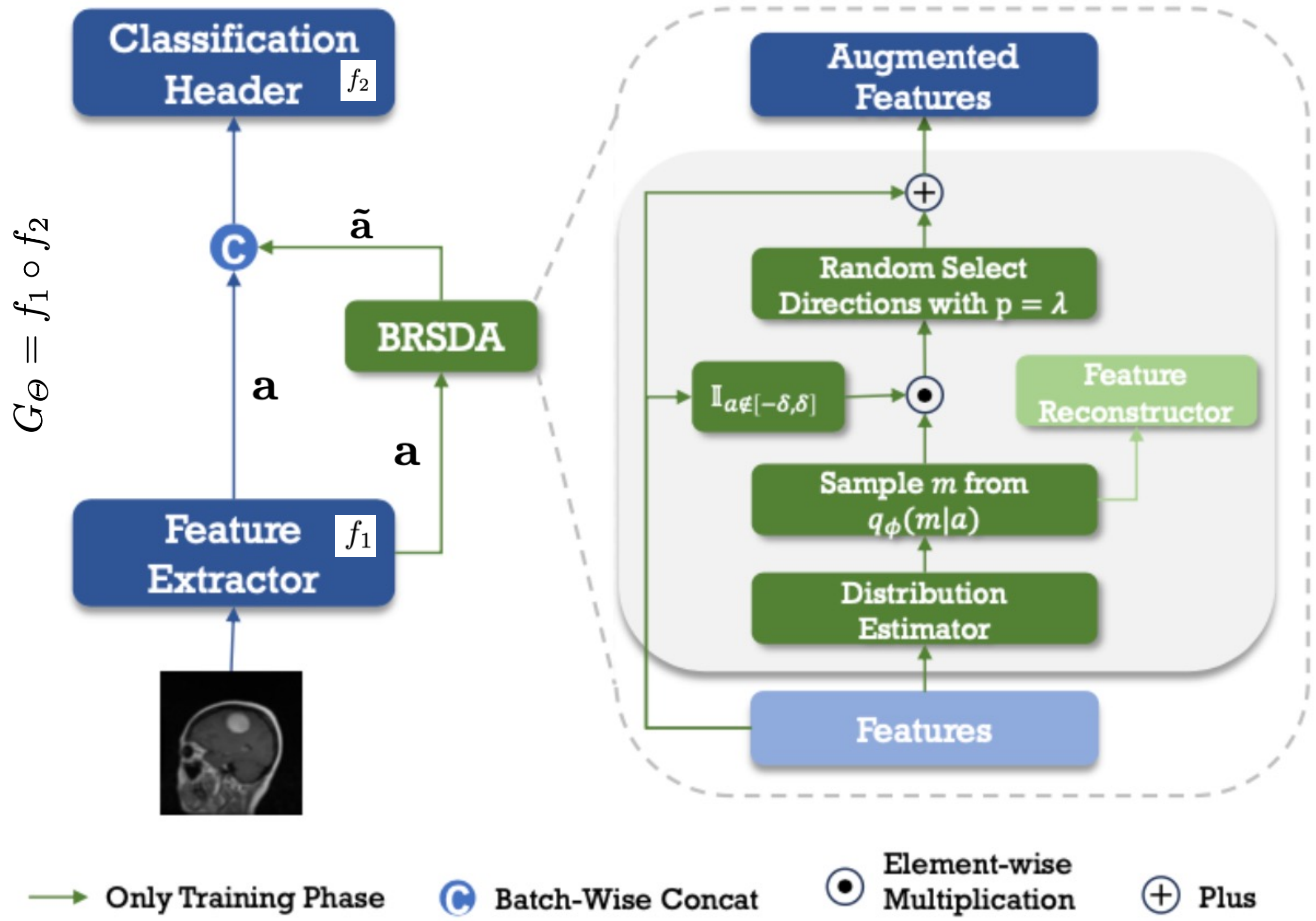


APPROACH – SETUP

$$G_\Theta = f_1 \circ f_2$$

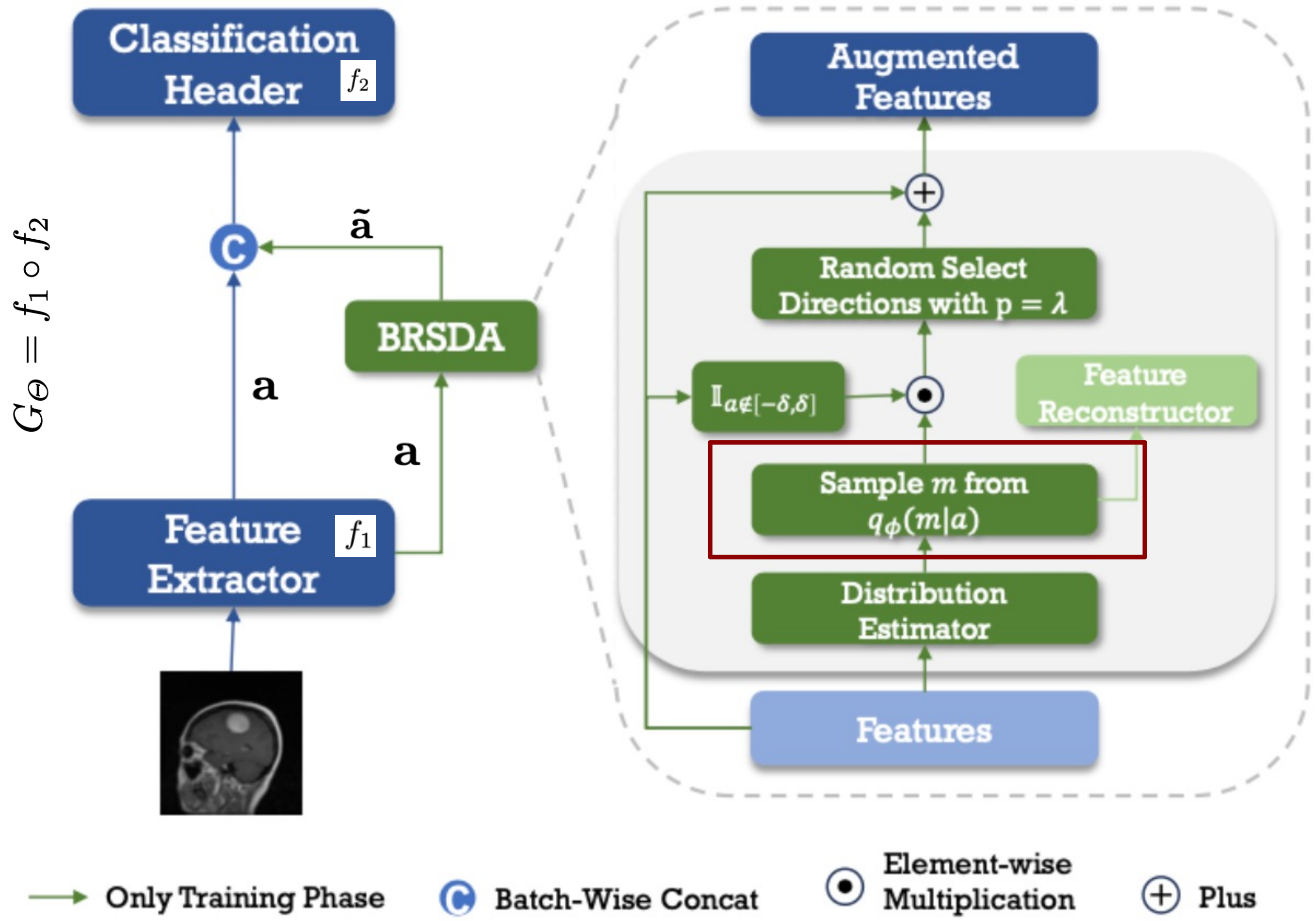


APPROACH – AUGMENTED FEATURE VECTOR



$$\tilde{\mathbf{a}} = \mathbf{a} + \mathbb{I}_{\mathbf{a} \neq 0} \mathbf{d}_\lambda \odot \mathbf{m}$$

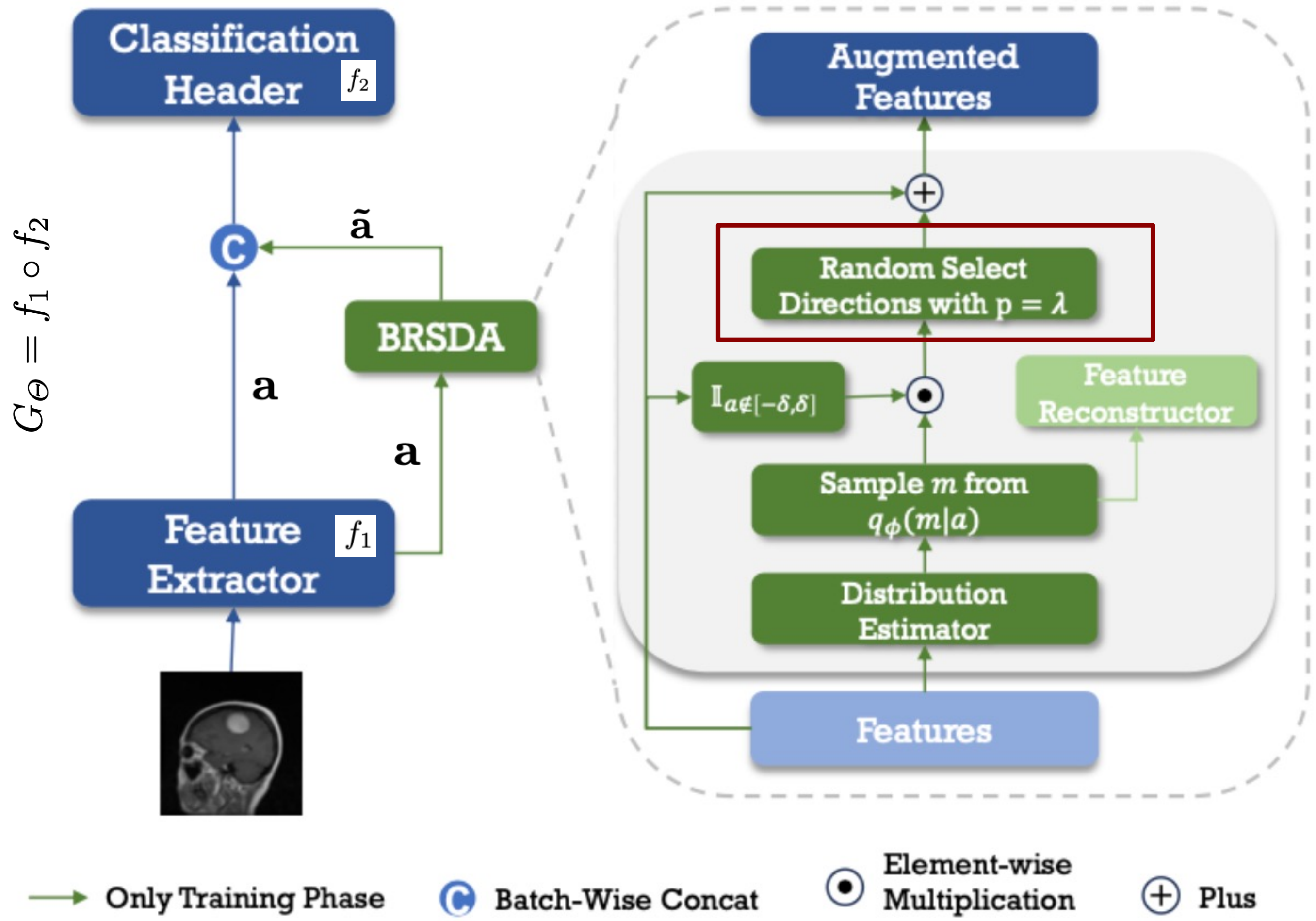
APPROACH – AUGMENTED FEATURE VECTOR



$$\tilde{\mathbf{a}} = \mathbf{a} + \mathbb{I}_{\mathbf{a} \neq 0} \mathbf{d}_\lambda \odot \mathbf{m}$$

- $\mathbf{m}$ refers to the semantic magnitude

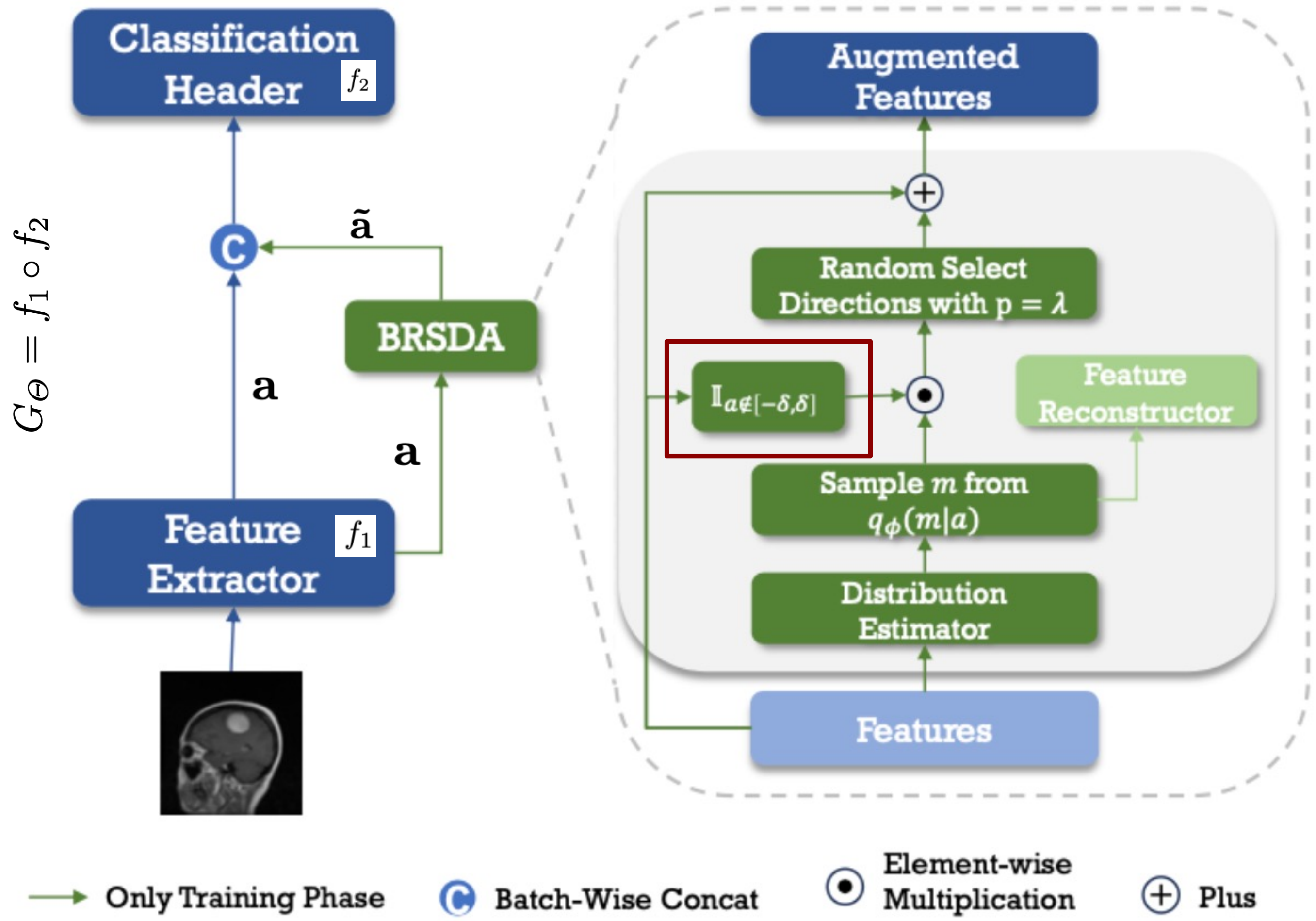
APPROACH – AUGMENTED FEATURE VECTOR



$$\tilde{\mathbf{a}} = \mathbf{a} + \mathbb{I}_{\mathbf{a} \neq 0} \mathbf{d}_\lambda \odot \mathbf{m}$$

- $\mathbf{m}$ refers to the semantic magnitude
- $\mathbf{d}_\lambda$ refers to the semantic directions

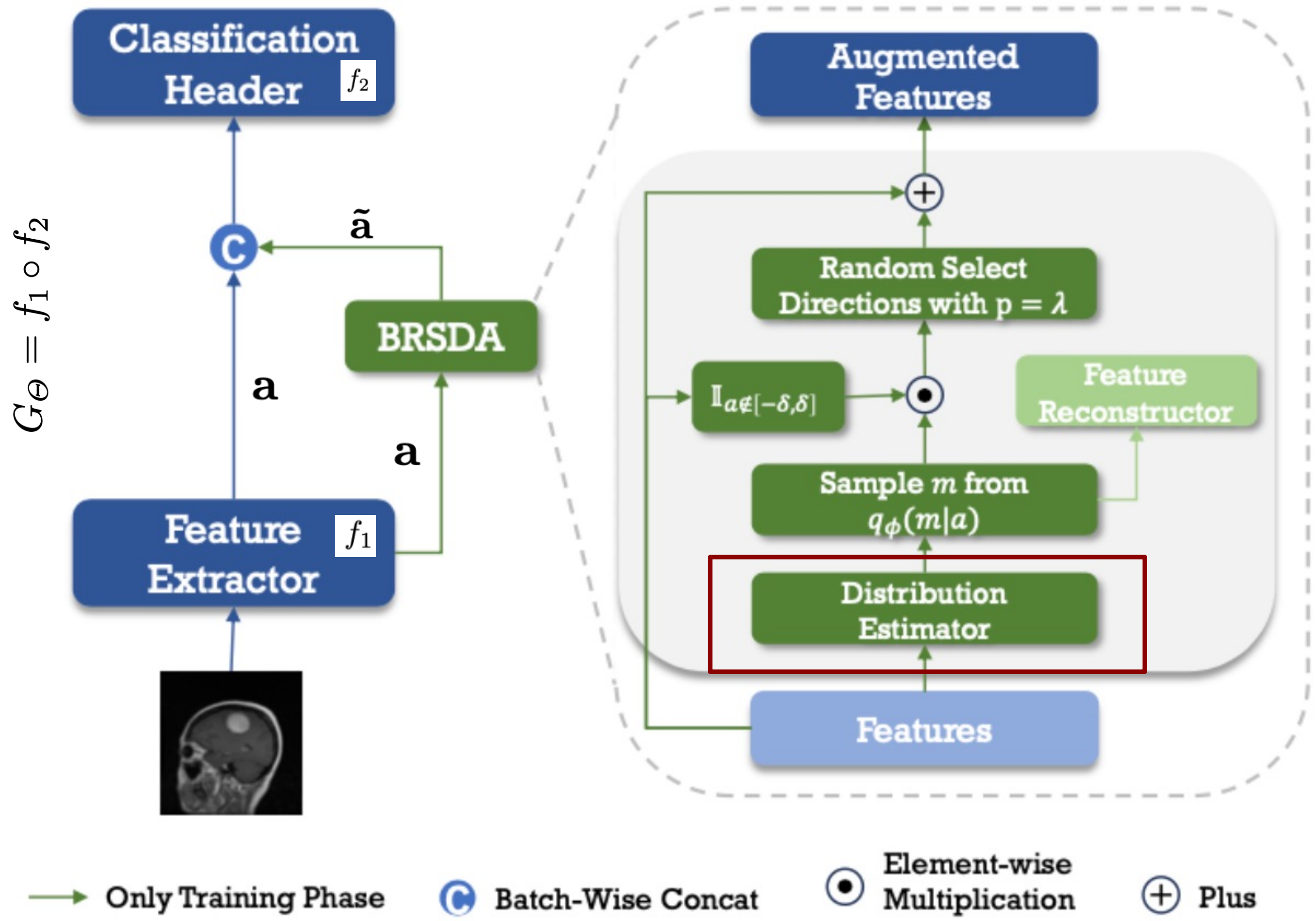
APPROACH – AUGMENTED FEATURE VECTOR



$$\tilde{\mathbf{a}} = \mathbf{a} + \mathbb{I}_{\mathbf{a} \neq 0} \mathbf{d}_\lambda \odot \mathbf{m}$$

- $\mathbf{m}$ refers to the semantic magnitude
- $\mathbf{d}_\lambda$ refers to the semantic directions
- $\mathbb{I}$ is an indicator function

APPROACH – SEMANTIC MAGNITUDE DISTRIBUTION



$$\tilde{\mathbf{a}} = \mathbf{a} + \mathbb{I}_{\mathbf{a} \neq 0} \mathbf{d}_\lambda \odot \mathbf{m}$$

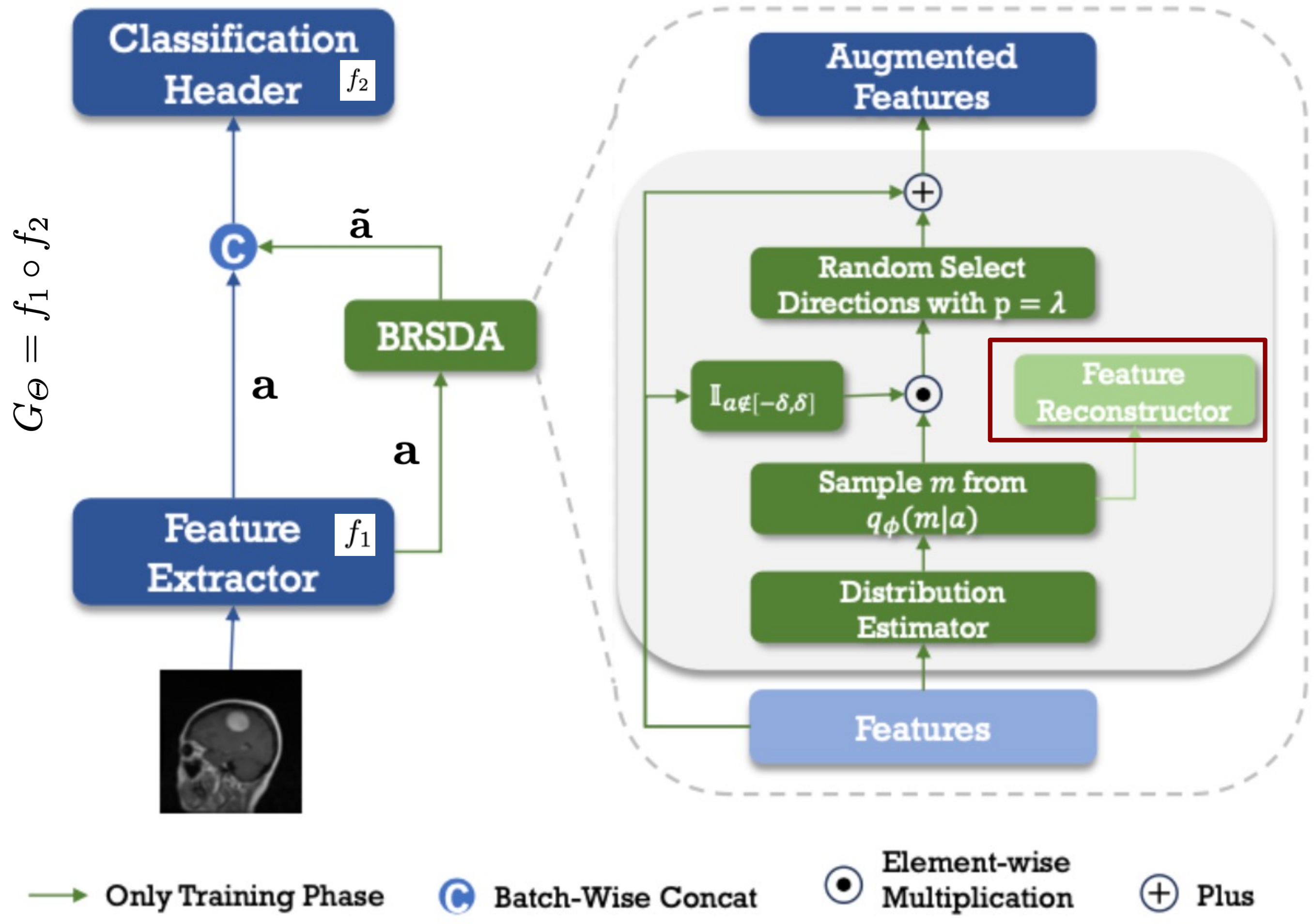
- $\mathbf{m}$ refers to the semantic magnitude
- We need to estimate the magnitude distribution
- Use model $q_{\phi_\mathbf{m}}(\mathbf{m}|\mathbf{a})$ approximate true distribution $p(\mathbf{m}|\mathbf{a})$
- Optimization with Kullback-Leibler divergence

$$\tilde{\phi}_\mathbf{m} = \arg\max_{\phi_\mathbf{m}} D_{KL}(q_{\phi_\mathbf{m}}(\mathbf{m}|\mathbf{a}) || p(\mathbf{m}|\mathbf{a})).$$

- Which expands to:

$$D_{KL}(q_{\phi_\mathbf{m}}(\mathbf{m}|\mathbf{a}) || p(\mathbf{m}|\mathbf{a})) = \mathrm{KL}(q_{\phi_\mathbf{m}}(\mathbf{m}|\mathbf{a}) || p(\mathbf{m})) - \mathbb{E}_{\mathbf{m} \sim q_{\phi_\mathbf{m}}(\mathbf{m}|\mathbf{a})}(\log p(\mathbf{a}|\mathbf{m})),$$

APPROACH – LOSS FUNCTION



$$D_{KL}(q_{\phi_\mathbf{m}}(\mathbf{m}|\mathbf{a})||p(\mathbf{m}|\mathbf{a})) = \mathrm{KL}(q_{\phi_\mathbf{m}}(\mathbf{m}|\mathbf{a})||p(\mathbf{m})) - \mathbb{E}_{\mathbf{m}\sim q_{\phi_\mathbf{m}}(\mathbf{m}|\mathbf{a})}(\log p(\mathbf{a}|\mathbf{m})).$$

- Second term is learned using a reconstruction network.
- BRSDA Loss Function:

$$\mathcal{L}_{brsda}(\phi_\mathbf{m}, \phi_\mathbf{a}; \mathbf{a}) = -\mathrm{KL}(q_{\phi_\mathbf{m}}(\mathbf{m}|\mathbf{a})||p(\mathbf{m})) + \mathbb{E}_{\mathbf{m}\sim q_{\phi_\mathbf{m}}(\mathbf{m}|\mathbf{a})}(\log p_{\phi_\mathbf{a}}(a|m)).$$

- Computed using reparameterization trick
- Total network loss

$$\mathcal{L} = \mathcal{L}^{\mathbf{a}}_{task} + \alpha(\mathcal{L}_{brsda} + \mathcal{L}^{\tilde{\mathbf{a}}}_{task}).$$

EVALUATION

Dataset	AUC%		ACC%		Modality
	Baseline	**BRSDA**	Baseline	**BRSDA**	
BreastMNIST [26]	89.62	**92.13**$_{(\mathbf{2.51})}$	84.62	**87.82**$_{(\mathbf{3.21})}$	Breast Ultrasound
RetinaMNIST [26]	71.69	**72.60**$_{(\mathbf{0.91})}$	45.00	**52.50**$_{(\mathbf{7.50})}$	Fundus Camera
LUNG [20]	87.48	**89.05**$_{(\mathbf{1.58})}$	77.31	**78.72**$_{(\mathbf{1.42})}$	X-Ray
BTMRI [18]	99.81	**99.93**$_{(\mathbf{0.12})}$	**97.64**	97.48$_{(-0.15)}$	MRI
CATAR [1]	97.98	**98.88**$_{(\mathbf{0.90})}$	89.26	**92.56**$_{(\mathbf{3.31})}$	Camera
OrganMNIST3D [26]	99.35	**99.38**$_{(\mathbf{0.04})}$	89.02	**89.51**$_{(\mathbf{0.49})}$	CT
NoduleMNIST3D [26]	89.65	**91.31**$_{(\mathbf{1.66})}$	**87.42**	85.81$_{(-1.61)}$	CT
AdernalMNIST3D [26]	88.60	**89.13**$_{(\mathbf{0.54})}$	71.81	**83.89**$_{(\mathbf{12.08})}$	Shape from CT
FractureMNIST3D [26]	74.09	**77.17**$_{(\mathbf{3.09})}$	52.50	52.50$_{(0.00)}$	CT
VesselMNIST3D [26]	95.79	**96.47**$_{(\mathbf{0.68})}$	94.50	94.50$_{(0.00)}$	Shape from MRI
SynapseMNIST3D [26]	71.18	**76.06**$_{(\mathbf{4.89})}$	75.00	**76.99**$_{(\mathbf{1.99})}$	Electron Microscope

FOR LESS THAN 4% EXTRA TRAINING TIME…

THANK YOU

PRESENTATION BY DIONYSIOS RIGATOS

ARXIV URL