Let $\mathbf{X} \in \mathbb{R}^{n\times p}$ denote a feature matrix, where n and p are respectively the numbers of samples (subjects) and variables (voxels). $\mathbf{y} \in \{0,1\}^n$ denote the class (diagnosis) vector of all the samples.
The classical ridge regression is to estimate the coefficients $\mathbf{w} \in \mathbb{R}^p$ so that :

$$\min_{\mathbf{w}} \|\mathbf{X}\mathbf{w} - \mathbf{y}\|_2^2 + \alpha\|\mathbf{w}\|_2^2 \tag{1}$$

where $\alpha > 0$ is a parameter of the coefficient penalization that controls the amount of shrinkage.
The proposed approach integrates the prior within the penalization term of (1), then :

$$\min_{\mathbf{w}} \|\mathbf{X}\mathbf{w} - \mathbf{y}\|_2^2 + \alpha\|\mathbf{w} - \lambda\mathbf{w_{prior}}\|_2^2 \tag{2}$$

where $\mathbf{w_{prior}}$ is the coefficient matrix from the prior which has been already learned. By making $\mathbf{b} = \mathbf{w} - \lambda\mathbf{w_{prior}}$, and the ridge regression formulation will remain the same, as it will return to solve $\mathbf{b}$.

Adding spatial penalization

$$\min_{\mathbf{w}} \|\mathbf{X}\mathbf{w} - \mathbf{y}\|_2^2 + \alpha\|\mathbf{w} - \lambda\mathbf{w_{prior}}\|_2^2 + \beta J(\mathbf{w}) \tag{3}$$

where $J(\mathbf{w})$ is the regularization :

$$J(\mathbf{w}) = \theta\|\mathbf{w}\|_{l_1} + (1-\theta)\|\mathbf{w}\|_{TV} \tag{4}$$

We write the loss function $\mathcal{L}$ as :

$$\begin{aligned}
\mathcal{L}(\mathbf{X},\mathbf{w},\mathbf{y}) &= \frac{1}{2}\|\mathbf{X}\mathbf{w} - \mathbf{y}\|_2^2 + \frac{\alpha}{2}\|\mathbf{w} - \lambda\mathbf{w_{prior}}\|_2^2 \\
&= \frac{1}{2}\mathbf{w^t X^T X w} + \frac{\alpha}{2}\mathbf{w^t w} - \mathbf{w^t X^T y} - \alpha\lambda\mathbf{w^t w_{prior}} + \frac{\alpha}{2}\lambda\mathbf{w^t_{prior} w_{prior}} + \frac{1}{2}\mathbf{y^t y} \\
&= \frac{1}{2}\mathbf{w^t}(\mathbf{X^T X} + \alpha\mathbf{I})\mathbf{w} - (\mathbf{X^T y} + \alpha\lambda\mathbf{w_{prior}})^{\mathbf{T}}\mathbf{w} + \mathbf{C} \\
&\equiv \frac{1}{2}\|\tilde{\mathbf{X}}\mathbf{w} - \tilde{\mathbf{y}}\|_{\mathbf{2}}^{\mathbf{2}}
\end{aligned}$$

where :

$$\tilde{\mathbf{X}} = \sqrt{\mathbf{X^T X} + \alpha\mathbf{I}} \tag{5}$$

and

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}}^{-\mathbf{1}}(\mathbf{X^T y} + \alpha\lambda\mathbf{w_{prior}}) \tag{6}$$