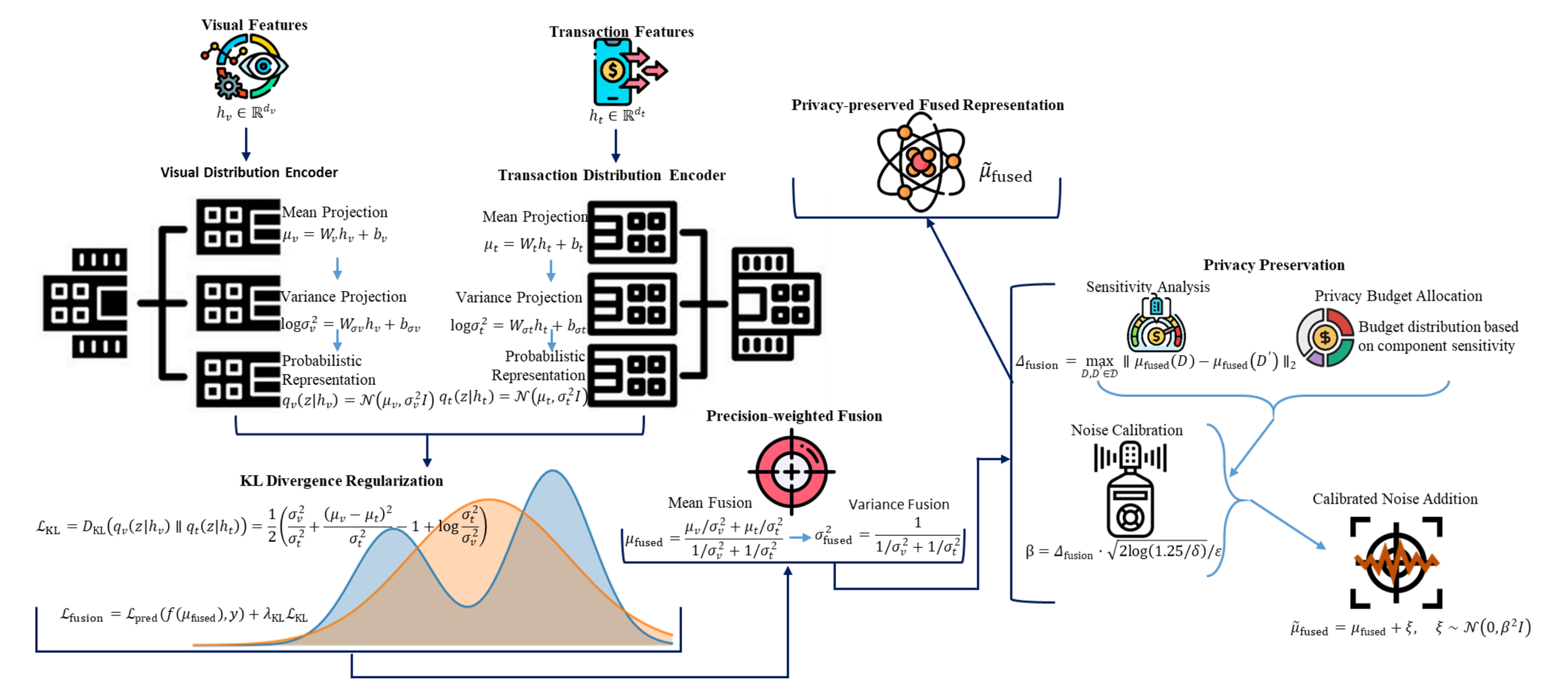
Visual Features
$h_v \in \mathbb{R}^{d_v}$
Transaction Features
$h_t \in \mathbb{R}^{d_t}$
Visual Distribution Encoder
Transaction Distribution Encoder
Mean Projection
$\mu_v = W_v h_v + b_v$
Variance Projection
$\log\sigma_v^2 = W_{\sigma v} h_v + b_{\sigma v}$
Probabilistic Representation
$q_v(z|h_v) = \mathcal{N}(\mu_v, \sigma_v^2 I)$
Mean Projection
$\mu_t = W_t h_t + b_t$
Variance Projection
$\log\sigma_t^2 = W_{\sigma t} h_t + b_{\sigma t}$
Probabilistic Representation
$q_t(z|h_t) = \mathcal{N}(\mu_t, \sigma_t^2 I)$
Privacy-preserved Fused Representation
$\tilde{\mu}_{\text{fused}}$
Privacy Preservation
Sensitivity Analysis
Privacy Budget Allocation
Budget distribution based on component sensitivity
$\Delta_{\text{fusion}} = \max_{D,D' \in \mathcal{D}} \| \mu_{\text{fused}}(D) - \mu_{\text{fused}}(D') \|_2$
Noise Calibration
$\beta = \Delta_{\text{fusion}} \cdot \sqrt{2\log(1.25/\delta)}/\varepsilon$
Calibrated Noise Addition
$\tilde{\mu}_{\text{fused}} = \mu_{\text{fused}} + \xi, \quad \xi \sim \mathcal{N}(0, \beta^2 I)$
Precision-weighted Fusion
Mean Fusion
$\mu_{\text{fused}} = \frac{\mu_v/\sigma_v^2 + \mu_t/\sigma_t^2}{1/\sigma_v^2 + 1/\sigma_t^2}$
Variance Fusion
$\sigma_{\text{fused}}^2 = \frac{1}{1/\sigma_v^2 + 1/\sigma_t^2}$
KL Divergence Regularization
$\mathcal{L}_{\text{KL}} = D_{\text{KL}}(q_v(z|h_v) \| q_t(z|h_t)) = \frac{1}{2}\left(\frac{\sigma_v^2}{\sigma_t^2} + \frac{(\mu_v - \mu_t)^2}{\sigma_t^2} - 1 + \log\frac{\sigma_t^2}{\sigma_v^2}\right)$
$\mathcal{L}_{\text{fusion}} = \mathcal{L}_{\text{pred}}(f(\mu_{\text{fused}}), y) + \lambda_{\text{KL}}\mathcal{L}_{\text{KL}}$