STAPLE CAM

STAPLE CAM ALGORITHM

- Simultaneous truth and performance level estimation (STAPLE), takes a collection of segmentations of an image, and computes simultaneously a probabilistic estimate of the true segmentation and a measure of the performance level represented by each segmentation. The source of each segmentation in the collection may be an appropriately trained human rater or raters, or it may be an automated segmentation algorithm. STAPLE algorithm is formulated by using the expectation maximization (EM) algorithm. In the formulation of algorithm, the expert segmentation decision at each voxel is directly observable, the hidden true segmentation is a binary variable for each voxel, and the performance level, or quality, achieved by each segmentation is represented by sensitivity and specificity parameters.

- The complete data consists of the segmentation decisions at each voxel, which are known, and the true segmentation, which is not known. If we did know the true segmentation, it would be straightforward to estimate the performance parameters by maximum-likelihood (ML) estimation. Since the complete data is not available, the complete data log likelihood cannot be constructed and instead must be estimated. Doing so requires evaluating the conditional probability density of the hidden true segmentation given the segmentation decisions and a previous estimate of the performance level of each segmentation generator. The expectation of the complete data log likelihood with respect to this density is then calculated, and from this estimate of the complete data log likehood, the performance parameters are found by ML estimation (or maximum a posteriori (MAP) estimation when a prior distribution for the parameters is considered). We iterate this sequence of estimation of the conditional probability of the true segmentation and performance parameters until convergence is reached. Convergence to a local maximum is guaranteed. Since we obtain both an estimate of the true segmentation and performance parameters from a collection of segmentations and a prior model, the algorithm is straightforward to apply to clinical imaging data. The algorithm enables the assessment of the performance of an automated image segmentation algorithm, and provides a simple method for direct comparison of human and algorithm performance.

DESCRIPTION OF STAPLE ALGORITHM

STAPLE estimates the performance parameters, and a probabilistic estimate of the true segmentation, by iterated estimation. The first step of each iteration is estimation of the conditional probability of the true segmentation given the expert decisions and previous performance parameter estimates and the second step is updated estimation of the performance parameters.

EXPECTATION(E) STEP OF EM ALGORITHM RESULTS

$$a_i^{(k)} \equiv f(T_i=1)\prod_j f\left(D_{ij}\middle|T_i=1, p_j^{(k)}, q_j^{(k)}\right) = f(T_i=1)\prod_{j:D_{ij}=1} p_j^{(k)} \prod_{j:D_{ij}=0}\left(1-p_j^{(k)}\right)$$

$$b_i^{(k)} \equiv f(T_i=0)\prod_j f\left(D_{ij}\middle|T_i=0, p_j^{(k)}, q_j^{(k)}\right) = f(T_i=0)\prod_{j:D_{ij}=0} q_j^{(k)} \prod_{j:D_{ij}=1}\left(1-q_j^{(k)}\right)$$

- In above formula notations are as follows:

1. D be an N × R matrix describing the binary decisions made for each segmentation at each voxel of the image.
2. T be an indicator vector of N elements, representing the hidden binary true segmentation, where for each voxel the structure of interest is recorded as present (1) or absent (0)
3. f (Ti =1) is the prior probability of Ti =1 and vice versa for f(Ti =0)
4. j:Dij = 1 denotes the set of indexes for which the decision of rater j at voxel i (i.e., Dij) has the value 1. vice versa for j: Dij =0
5. The probability mass function of the complete data be f (D, T |p , q).
6. The parameters pj , qj ∈ [0,1] are assumed to be characteristic of the rater, and may be equal for different raters but in general are not.
7. pj represent the "true positive fraction" or sensitivity (relative frequency of Dij = 1 when Ti = 1), $p_j = \Pr(D_{ij}=1 \mid T_i=1)$

   qj represent the "true negative fraction" or specificity (relative frequency of Dij = 0 when Ti = 0) $q_j = \Pr(D_{ij}=0 \mid T_i=0)$.

- Based on these expressions, we can now write a compact expression for the conditional probability of the true segmentation at each voxel. Using the notation common for EM algorithms, we refer to this as the weight variable

$$W_i^{(k-1)} \equiv f\left(T_i=1 \middle| D_i, p^{(k-1)}, q^{(k-1)}\right) = \frac{a_i^{(k-1)}}{a_i^{(k-1)}+b_i^{(k-1)}}.$$

- The weight indicates the probability of the true segmentation at voxel i being equal to 1. It is a normalized product of the prior probability of Ti = 1 (f (Ti = 1)), the sensitivity of each of the experts that decided the true segmentation was one and (1—sensitivity) of each of the experts that decided the true segmentation was zero.

MAXIMIZATION(M) STEP OF EM ALGORITHM RESULTS

- Given the estimated weight variables , which represent the conditional probabilities of the true segmentation, we can find the values of the expert performance level parameters $\left(p_j^{(k)}, q_j^{(k)}\right)$ that maximize the conditional expectation of the complete data log likelihood function.

$$\left(p^{(k)}, q^{(k)}\right) = \arg\max_{p,q} \sum_j \sum_i E\left[\ln f(D_{ij} \,|\, T_i, p_j, q_j) \,\middle|\, D, p^{(k-1)}, q^{(k-1)}\right]$$

- Based on maximizing these equation, we get the equations to find and update the values of pj and qj at each iteration of EM algorithm in terms of weight variable of previous iterations. Equation to update pj and qj values are as follows :

$$p_j^{(k)} = \frac{\sum_{i:D_{ij}=1} W_i^{(k-1)}}{\sum_i W_i^{(k-1)}}$$

$$q_j^{(k)} = \frac{\sum_{i:D_{ij}=0} \left(1 - W_i^{(k-1)}\right)}{\sum_i \left(1 - W_i^{(k-1)}\right)}.$$

- All the notations used in above equations are same as used in previous equations.

$W_i^{(k-1)}$ is the weight variable of previous iteration whose value we find in E- Step of Algorithm

INITIALIZATION OF STAPLE ALGORITHM

- STAPLE is most conveniently initialized by providing starting estimates for the sensitivity (pj) and specificity (qj) parameters. In the absence of other information regarding the relative quality of the experts or the true segmentation, initializing the sensitivity and specificity parameters to the same value and equal across all raters is recommended. This is equivalent up to a scaling factor to forming an initial estimate through a voting rule (such as assigning initial probabilities of voxels based on the frequency of selection by experts). In the paper on STAPLE CAM which used synthetic data and clinical experiments, successful estimation results were obtained by initializing all of the sensitivity and specificity parameters to the same values, with values close to but not equal to 1. for example, selecting An alternative strategy for initialization, useful when such information is available, is to provide an initial true segmentation estimate. An interesting strategy for certain problems is to use a probabilistic atlas to provide an initial true segmentation. For example, one may use a probabilistic atlas of the brain to provide an initial true segmentation estimate for the tissues of the brain.

CONVERGENCE OF STAPLE ALGORITHM

- The EM algorithm is guaranteed to converge to a local optimum because the algorithm is guaranteed to increase the likelihood at each iteration. The STAPLE estimates both performance parameters and the true segmentation, and convergence may be detected by monitoring these. A simple and good measure of convergence is the rate of change of the sum of the true segmentation probability, obtained by summing the each voxel evaluated at each iteration k using the formula :

$$S_k=\sum_{i=1}^{N} W_i$$

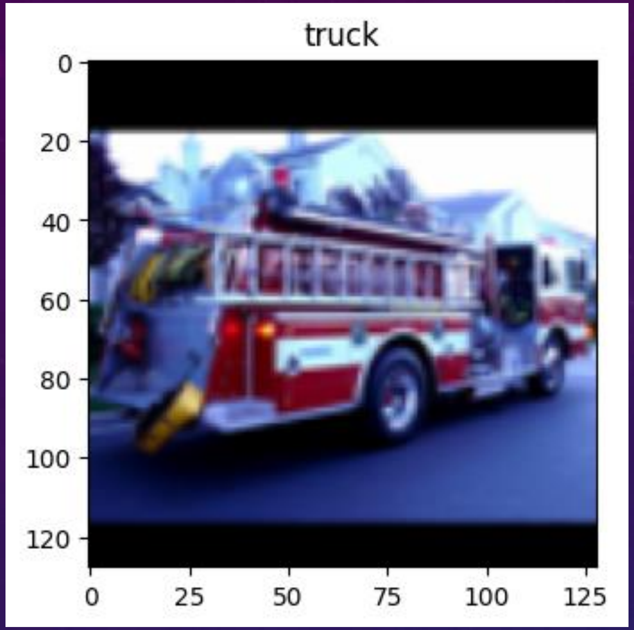
- We have found that iterating until $S_k - S_{k-1}$ = 0 or less than some preset threshold value, for example ε = 1 × 10−7. So once the value of the expression $S_k - S_{k-1}$ become less than these threshold value the algorithm is said to be converge, so it will stop the iteration of EM algorithm and we got final STAPLE segmented Image.

VISUALIZATION OF STAPLE ALGORITHM ON SAMPLE DATASETS

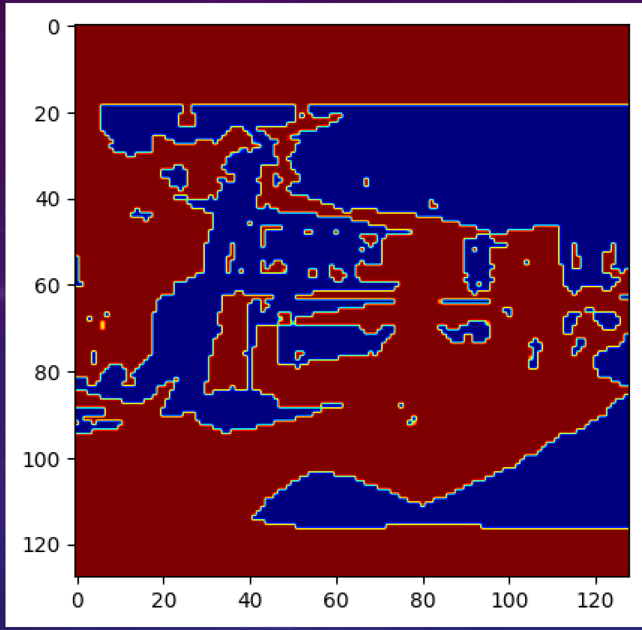
1. I have tested the STAPLE algorithm on 5 datasets : STL10, OxfordIIITPet, Flowers-102 and Caltech-101.
2. STL10 dataset is modified version of CIFAR10 dataset having images like cat, dog, deer, ship truck, etc. of different 10 classes.
3. OxfordIIITPet is dataset having images of animals from 37 different classes.
4. Flowers-102 is dataset having images of flowers from 102 different classes.

STL10 DATASET

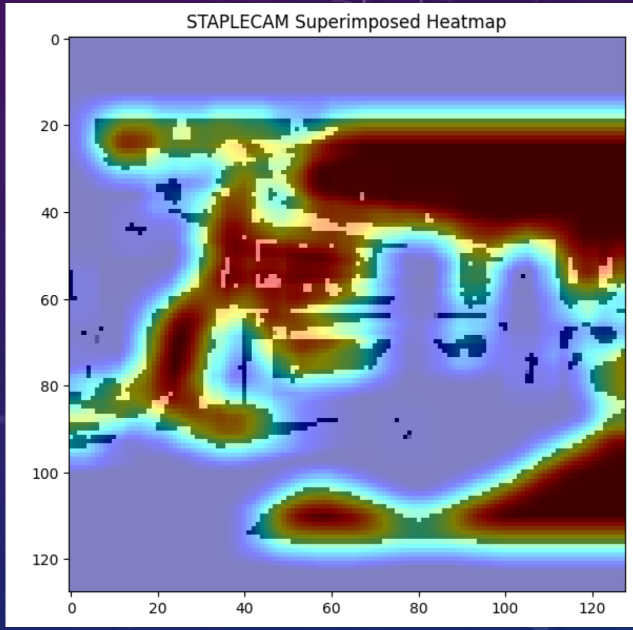
INPUT IMAGE

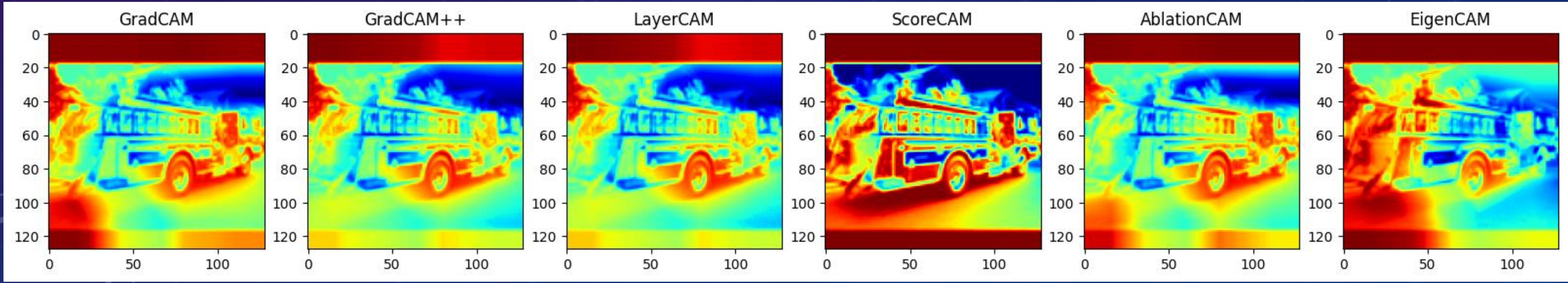


STAPLE CAM IMAGE



STAPLECAM SUPERIMPOSED HEATMAP

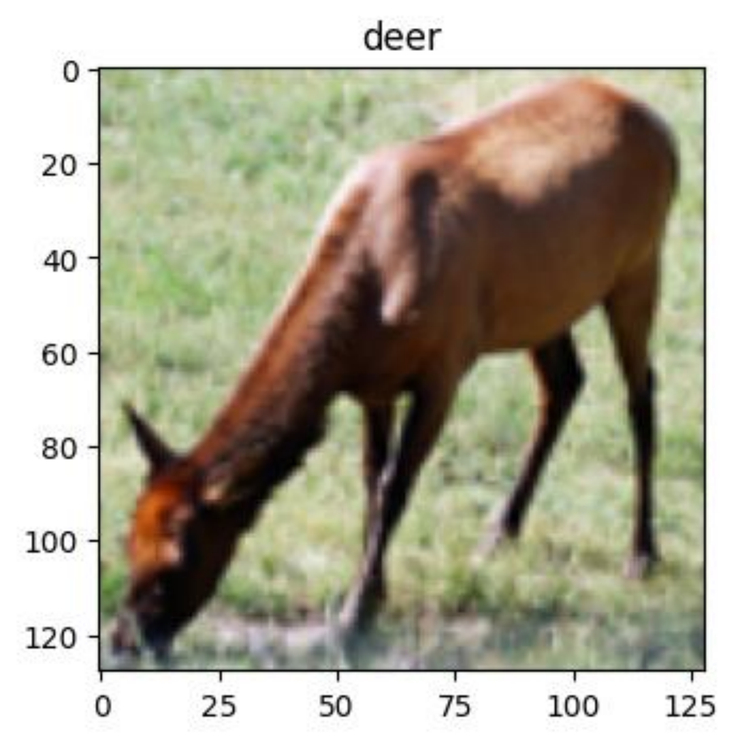


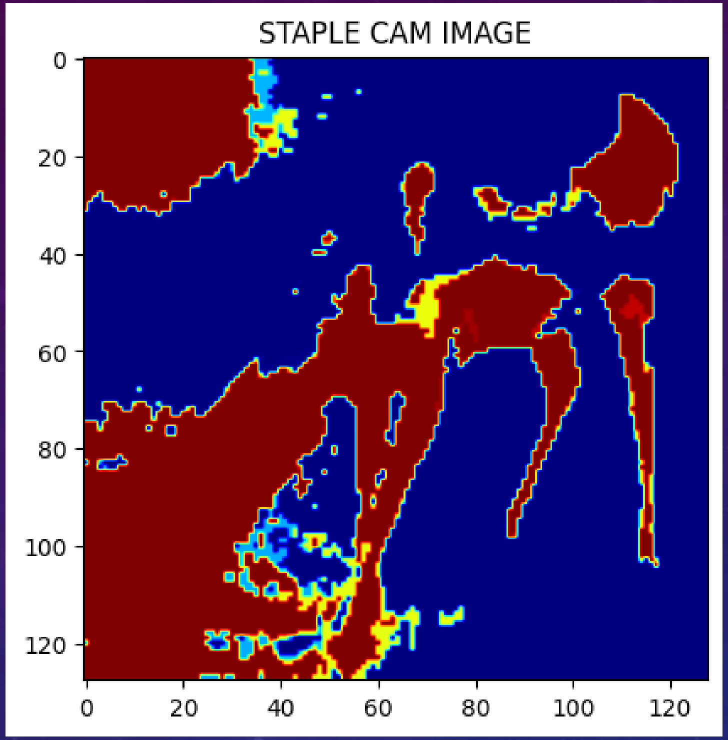


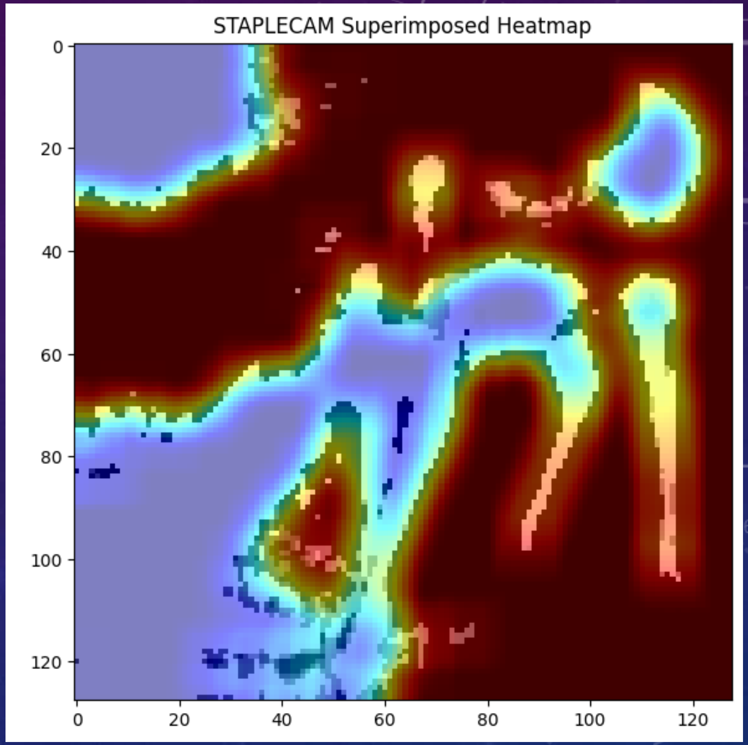
INPUT IMAGE

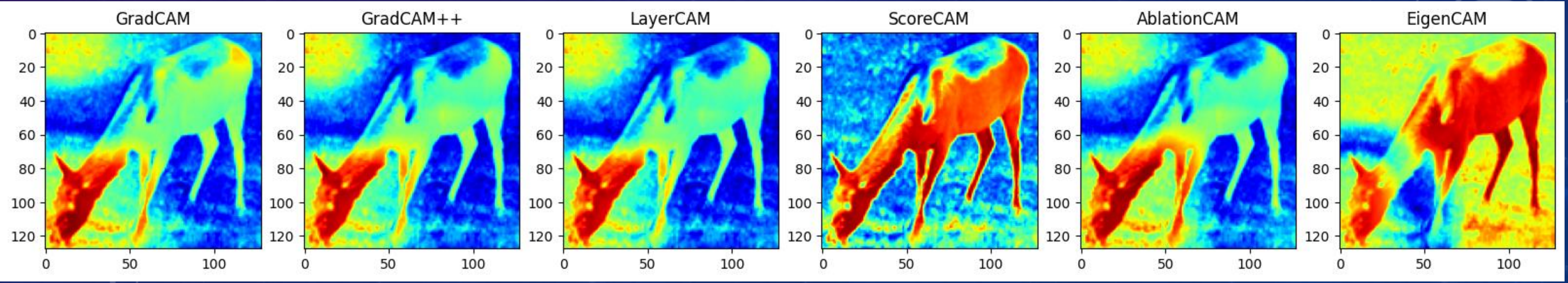
STAPLE CAM IMAGE

STAPLECAM SUPERIMPOSED HEATMAP

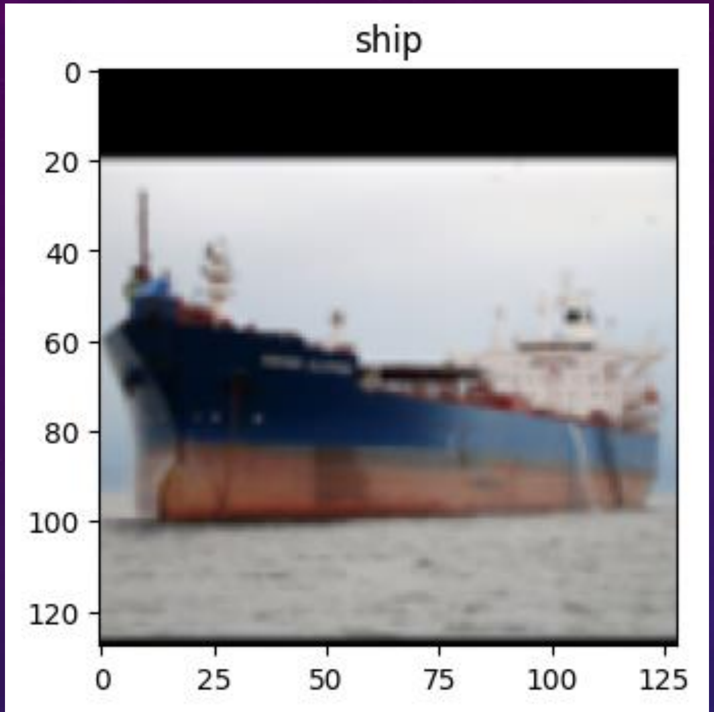




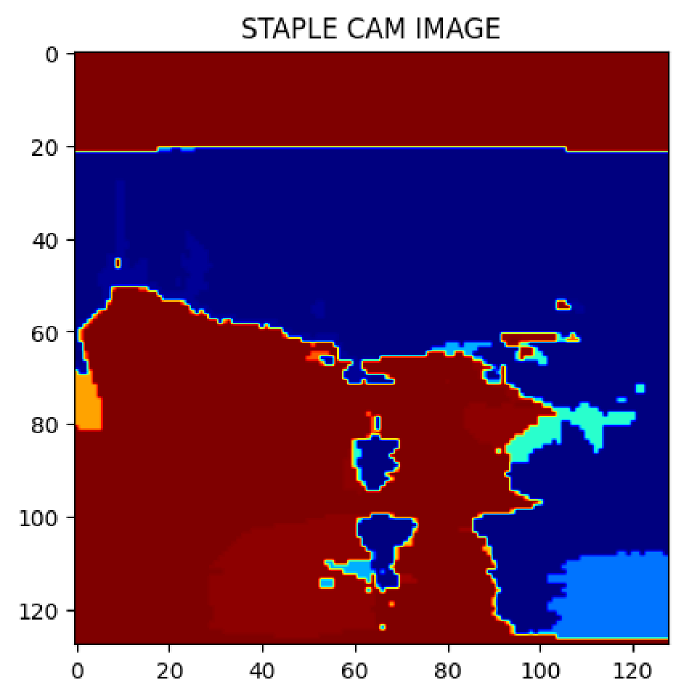




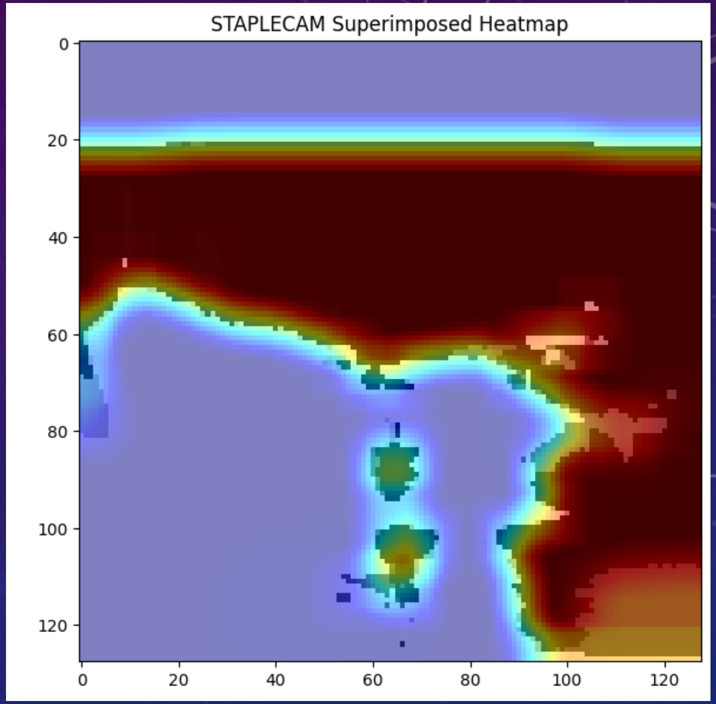
INPUT IMAGE

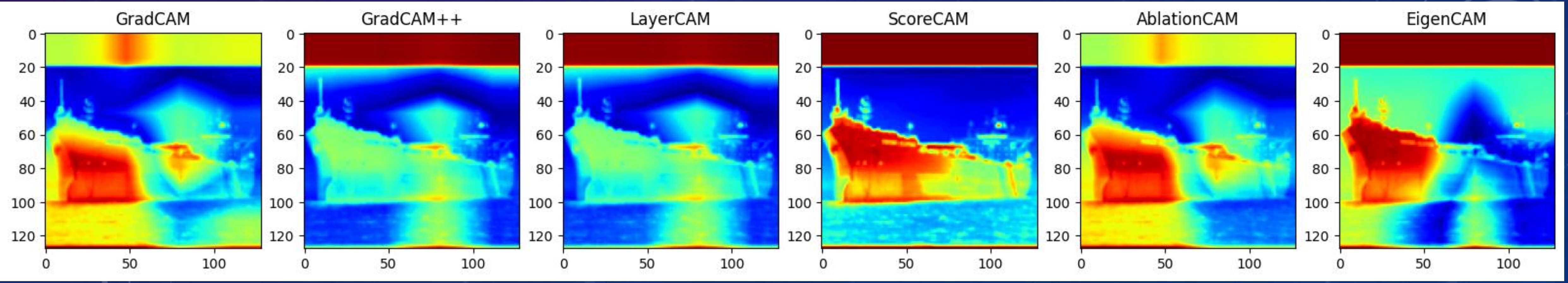


STAPLE CAM IMAGE

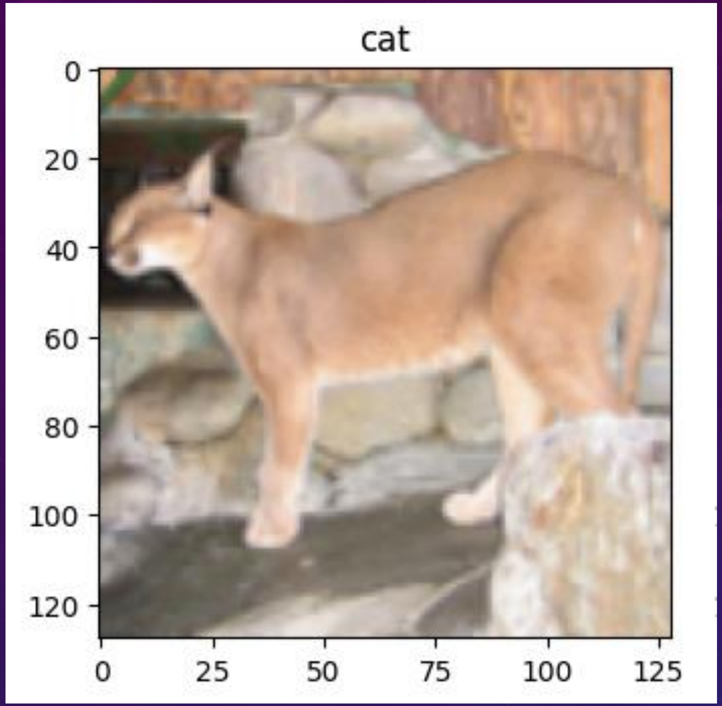


STAPLECAM SUPERIMPOSED HEATMAP

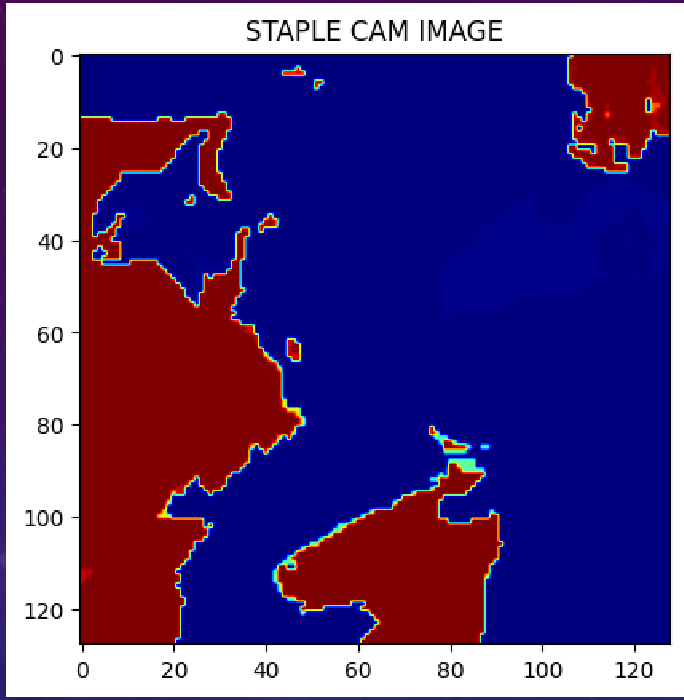




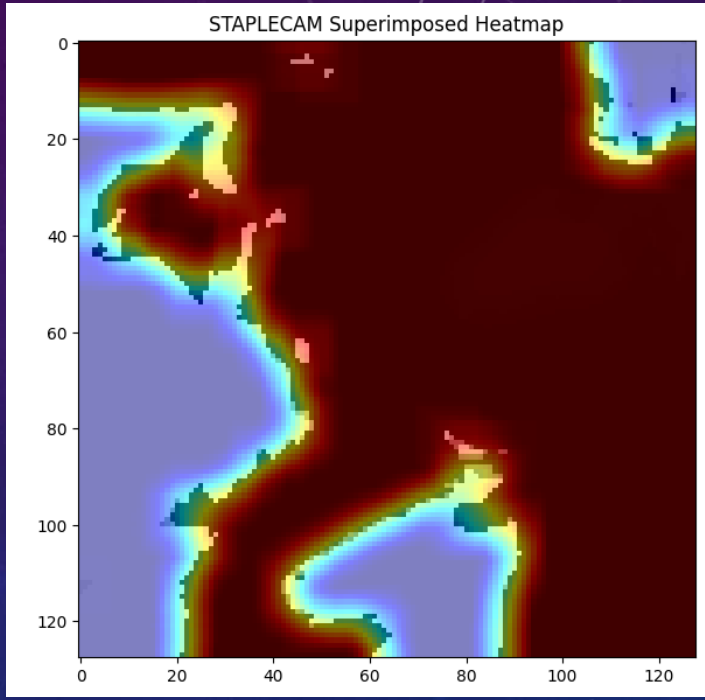
INPUT IMAGE

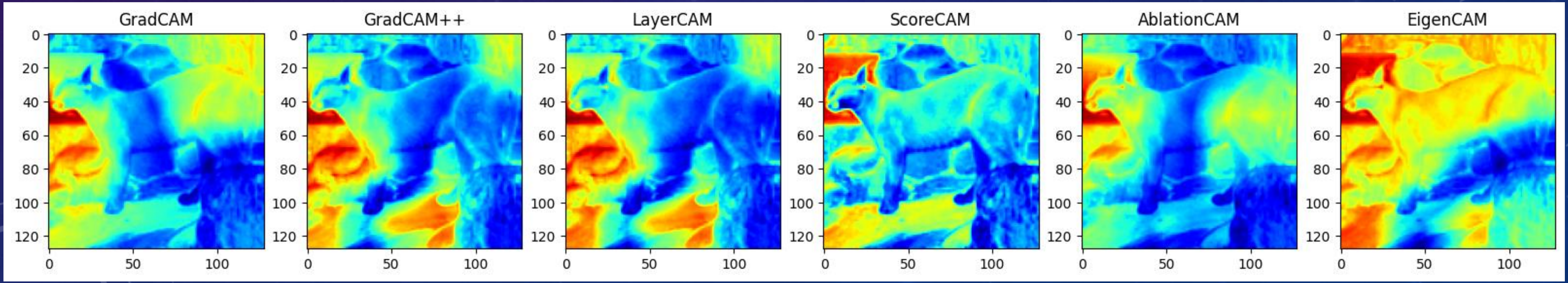


STAPLE CAM IMAGE

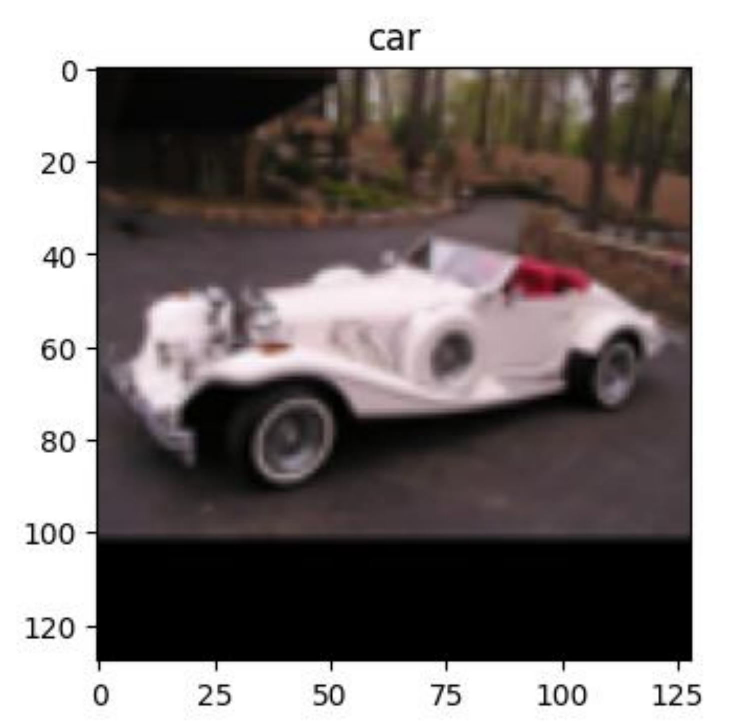


STAPLECAM SUPERIMPOSED HEATMAP

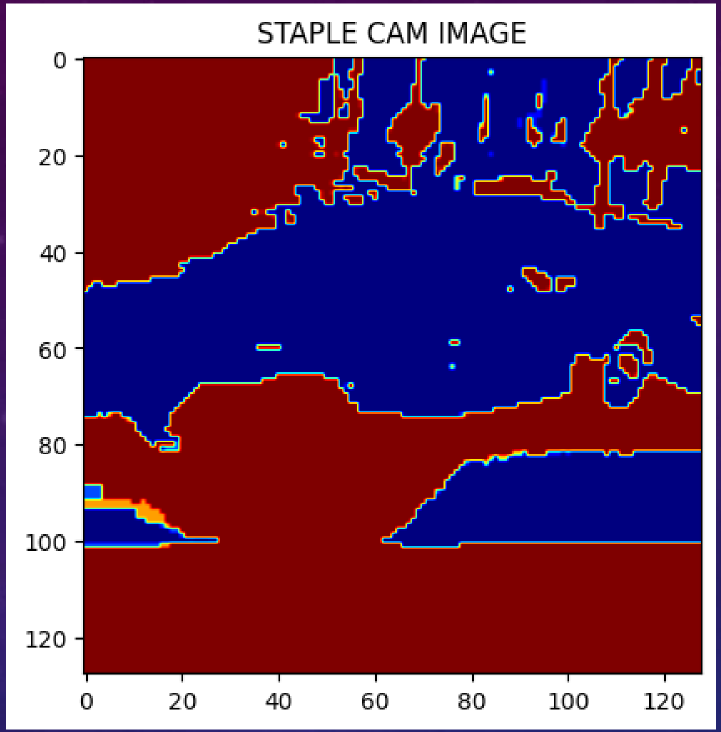




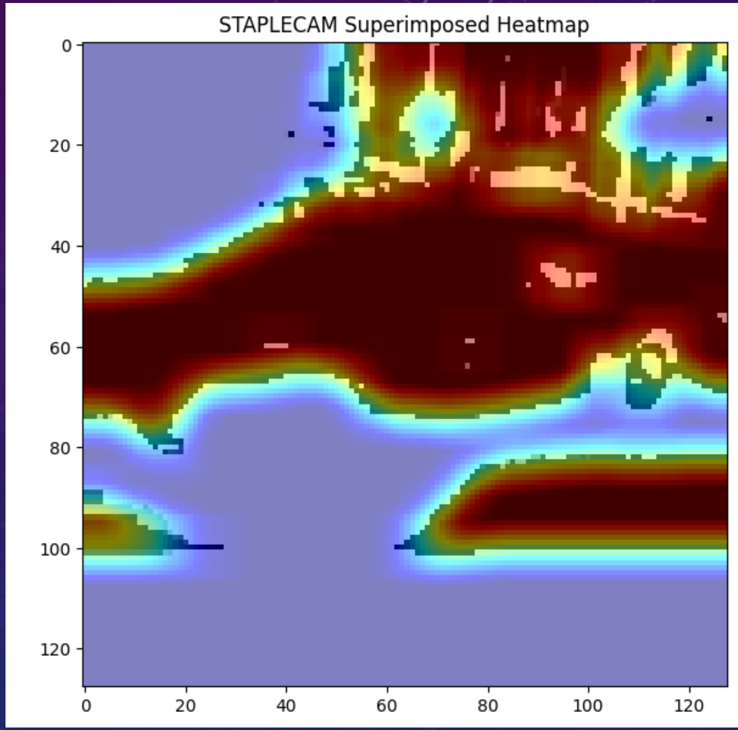
INPUT IMAGE

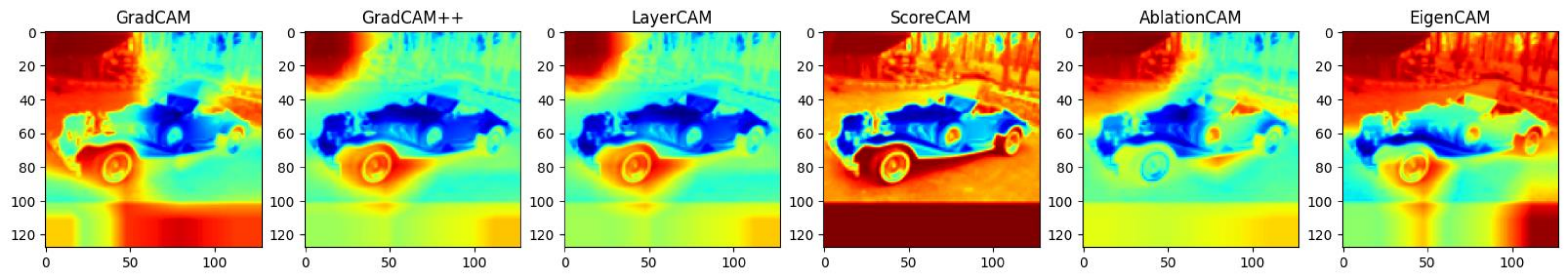


STAPLE CAM IMAGE

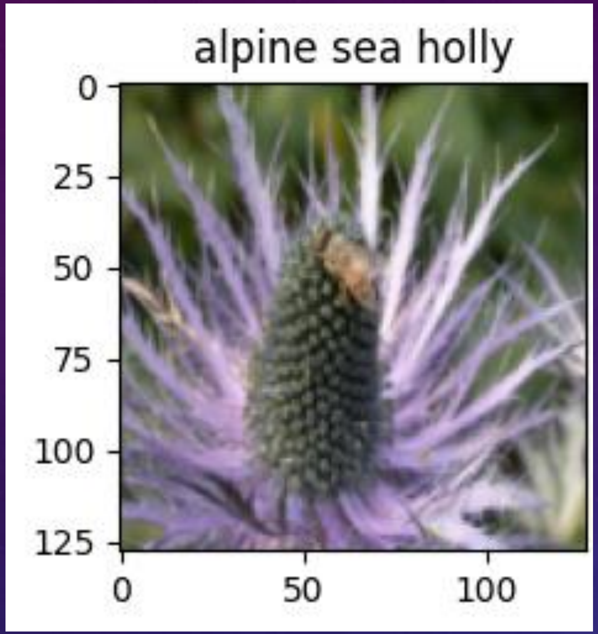


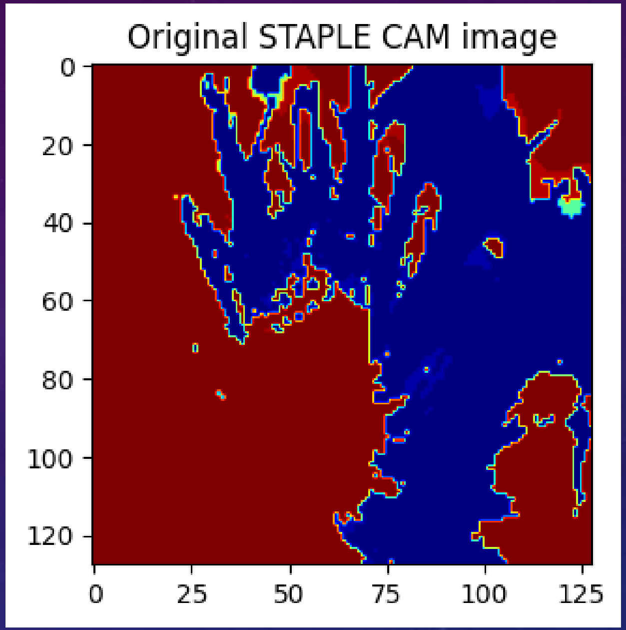
STAPLECAM SUPERIMPOSED HEATMAP

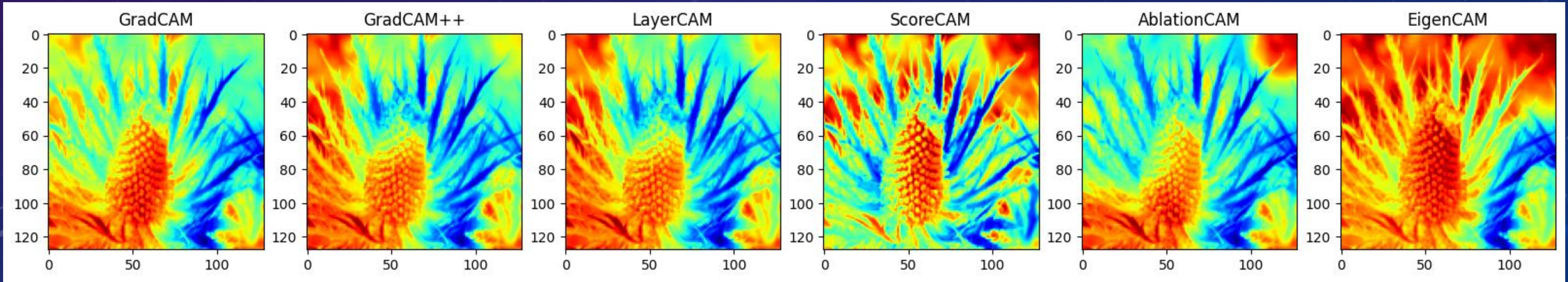




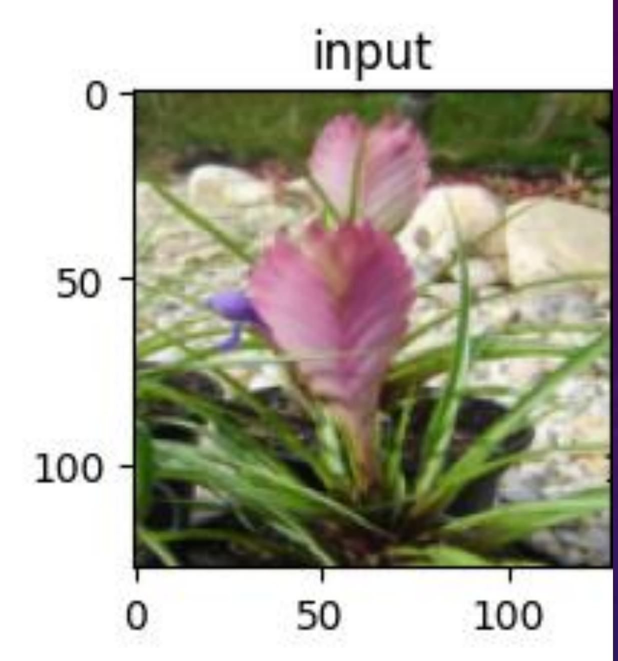
FLOWERS-102 DATASET



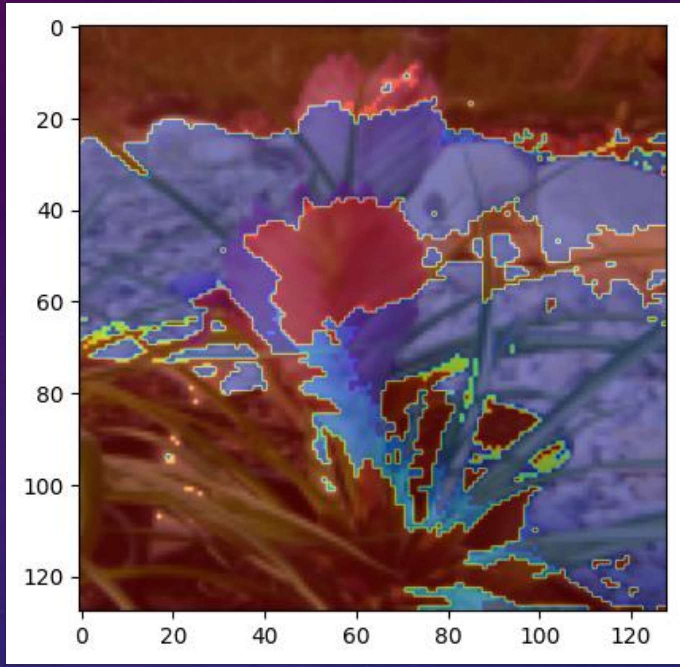




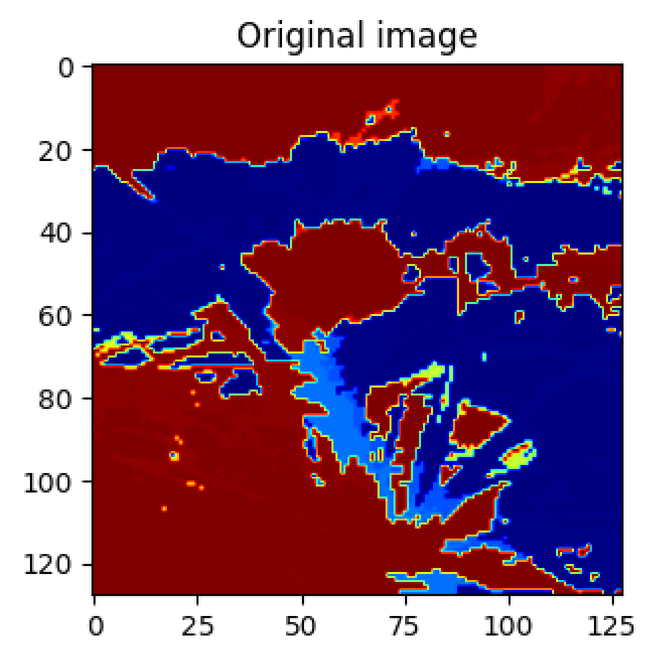
FOXGLOVE

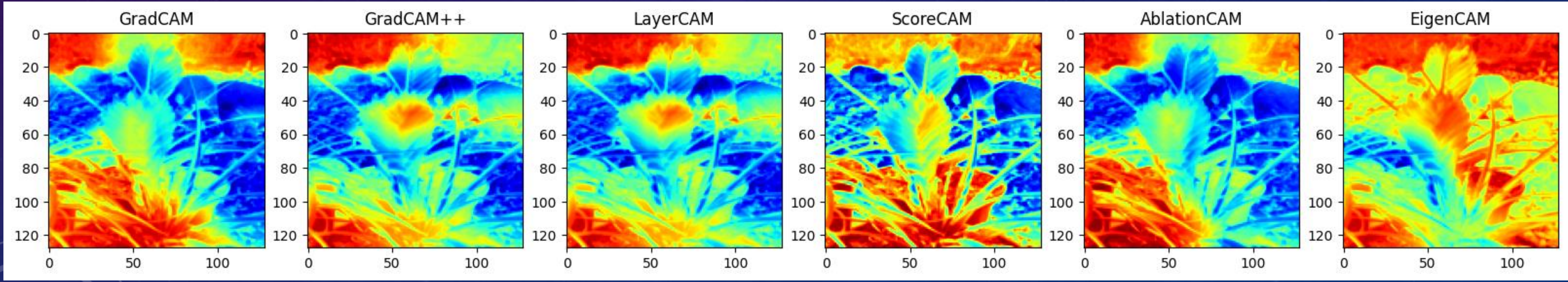


STAPLE CAM SUPERIMPOSED

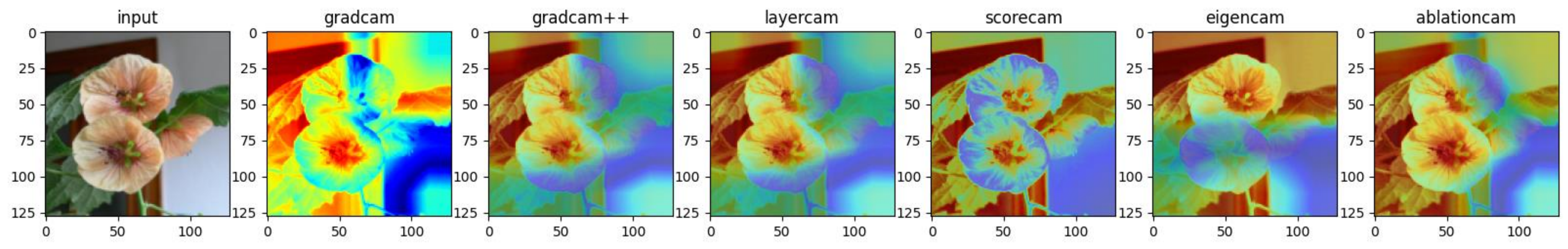


STAPLE CAM IMAGE

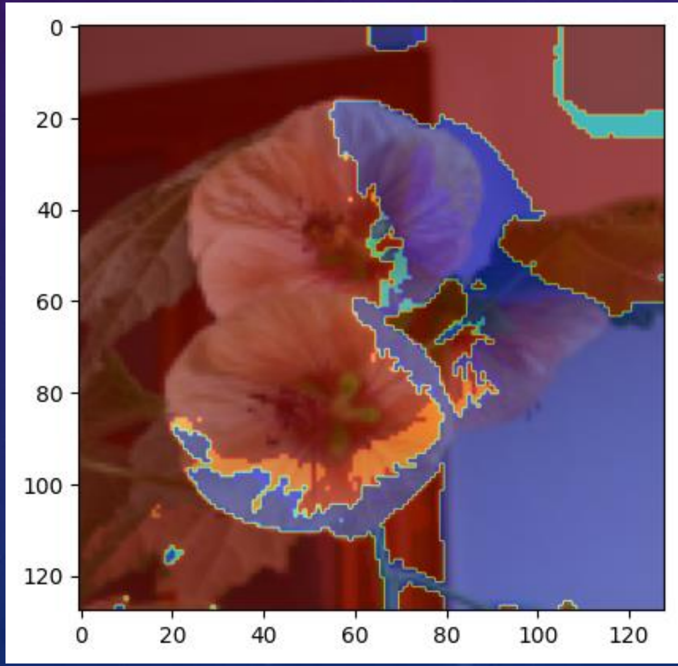




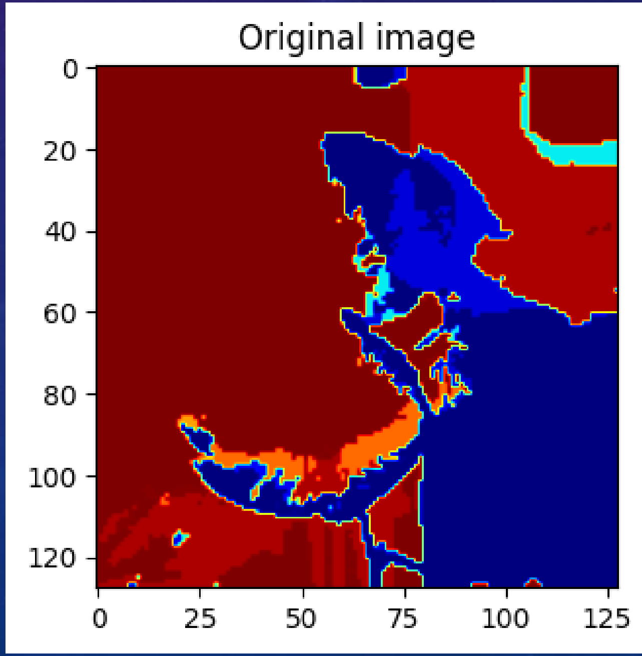
MEXICAN PETUNIA



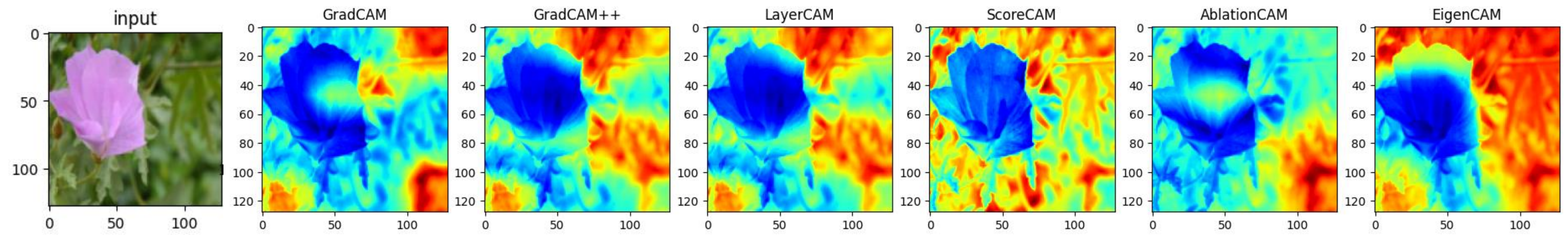
STAPLE CAM SUPERIMPOSED



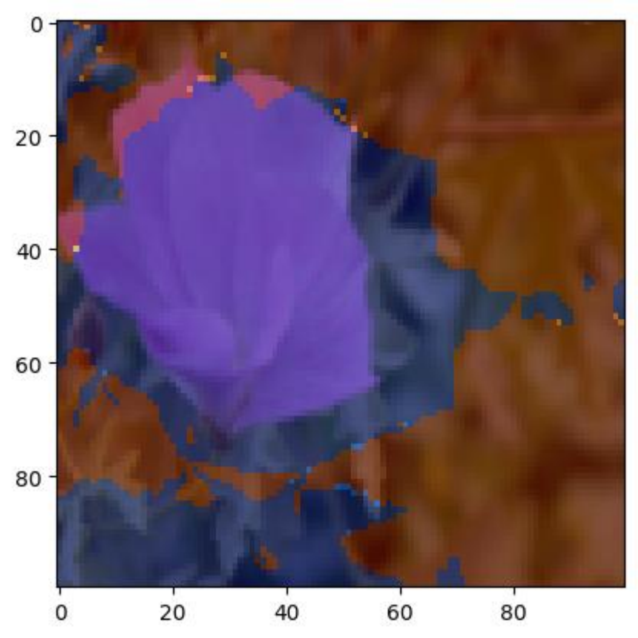
STAPLE CAM IMAGE



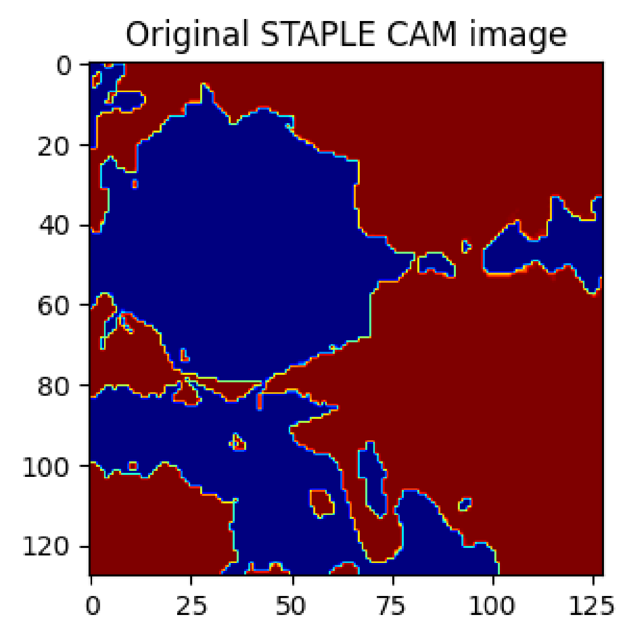
PELARGONIUM



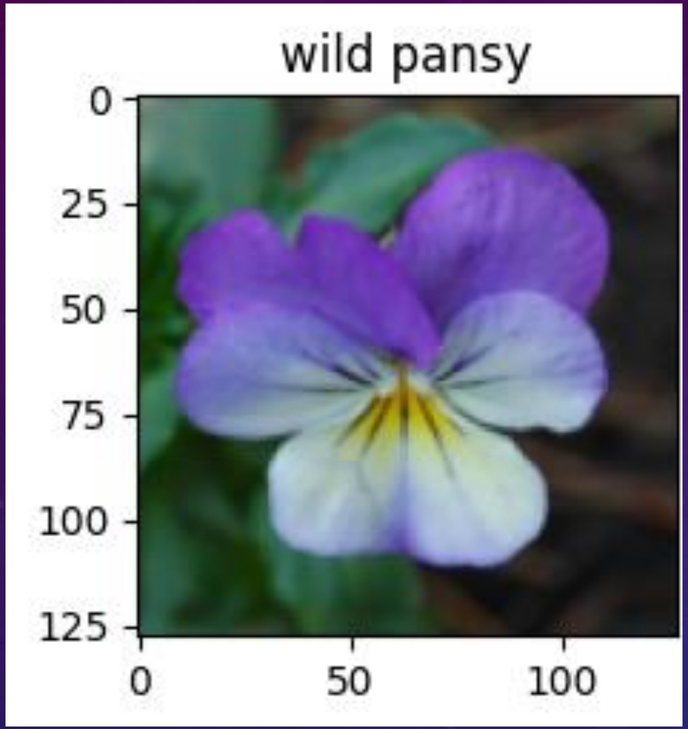
STAPLE CAM SUPERIMPOSED



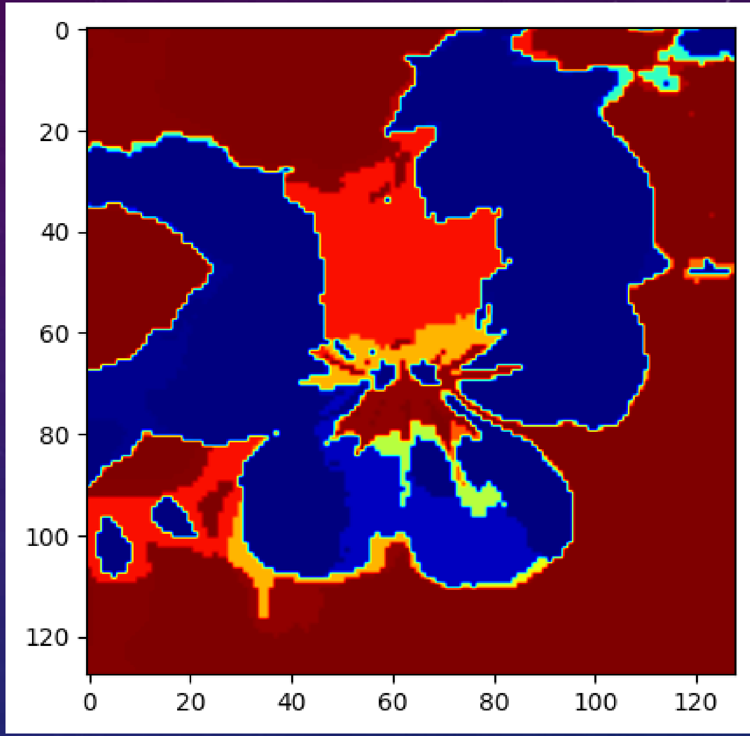
STAPLE CAM IMAGE

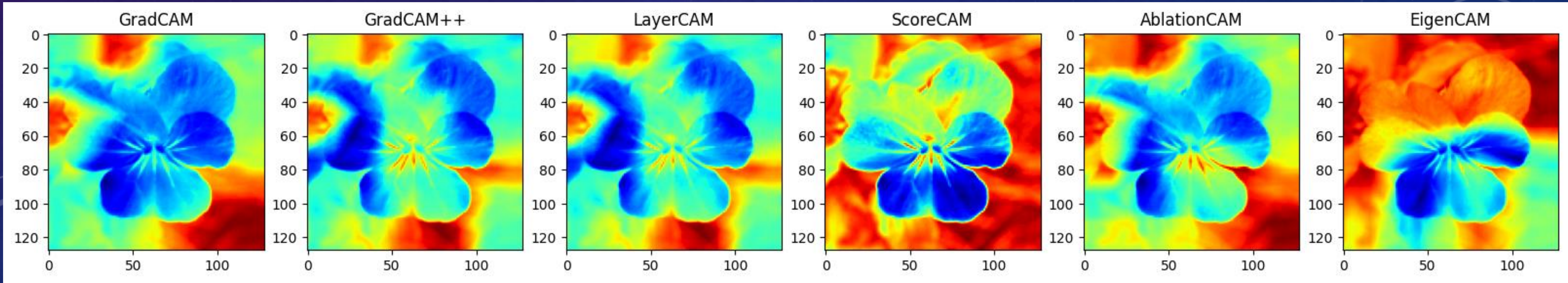


INPUT IMAGE



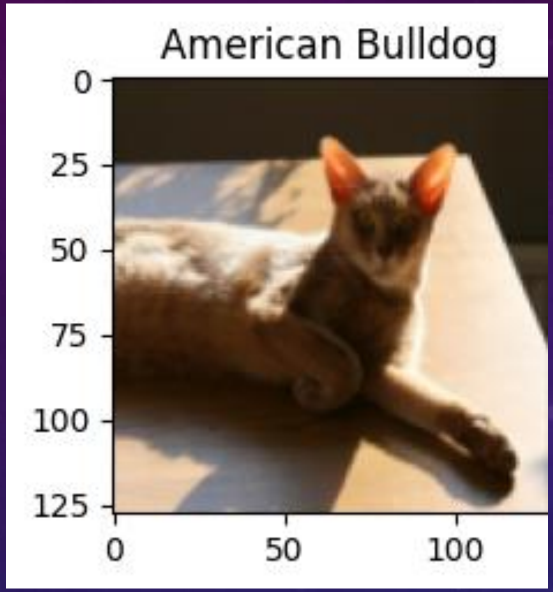
STAPLE CAM IMAGE



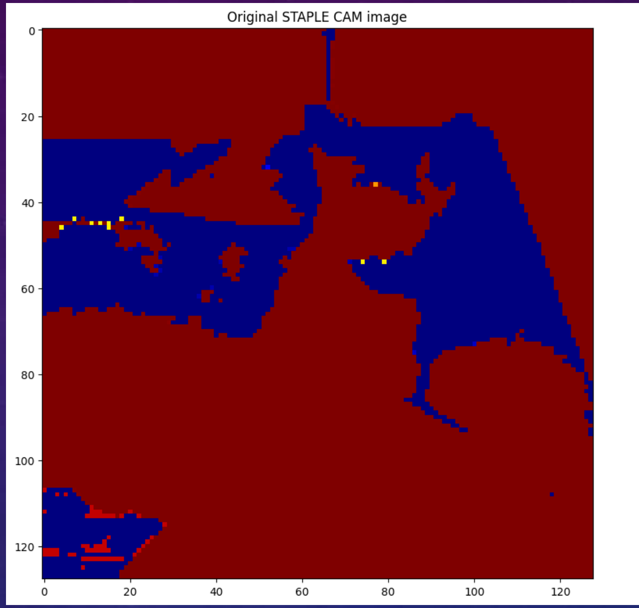


OXFORD IIITPET DATASET

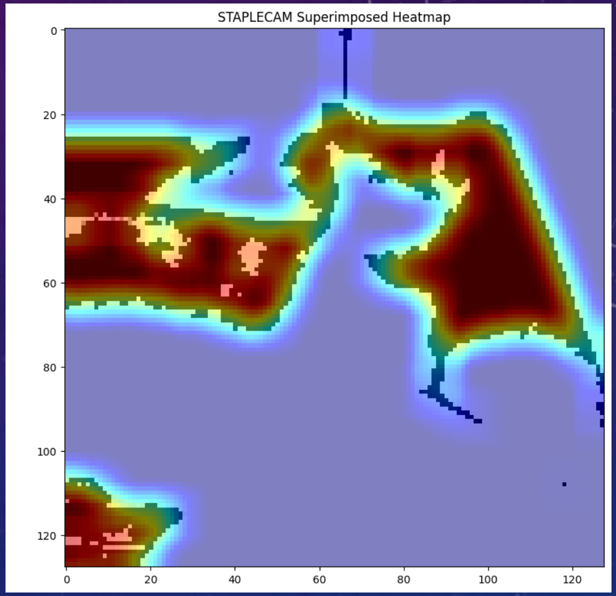
INPUT IMAGE

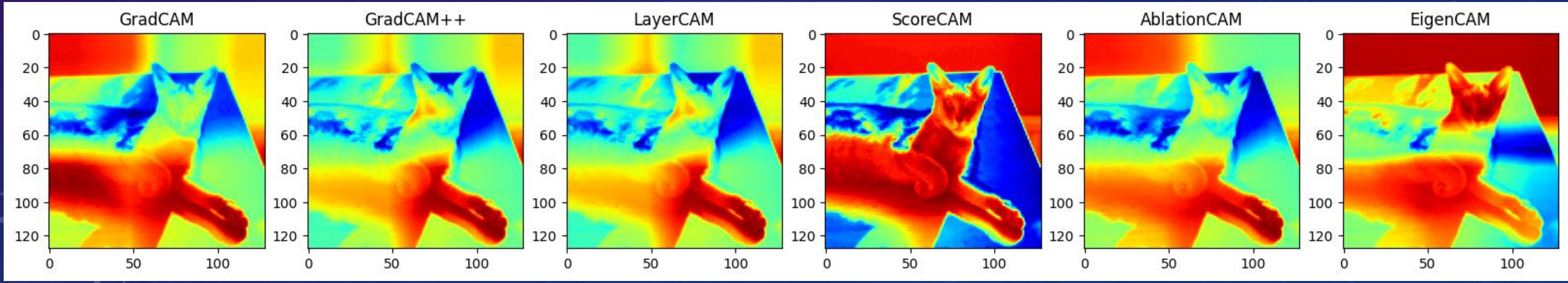


STAPLE CAM IMAGE



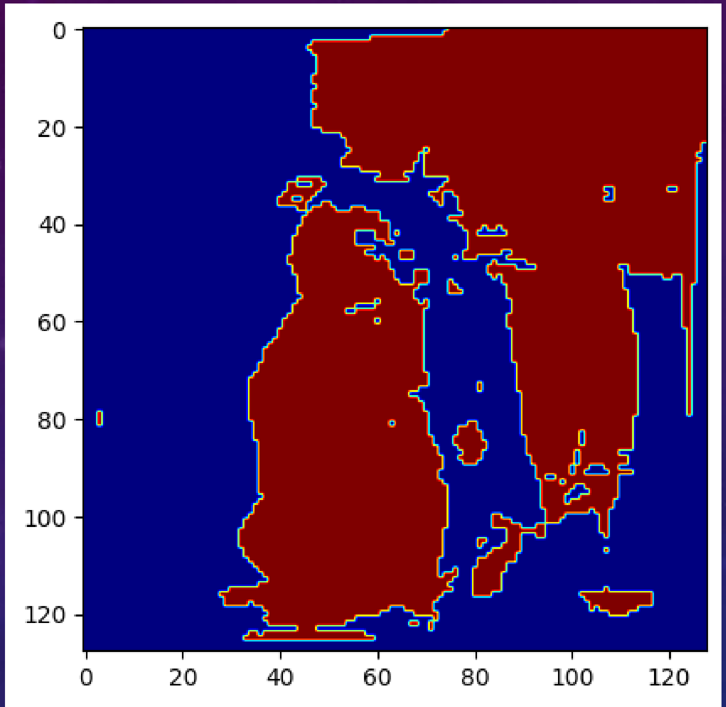
STAPLECAM SUPERIMPOSED HEATMAP



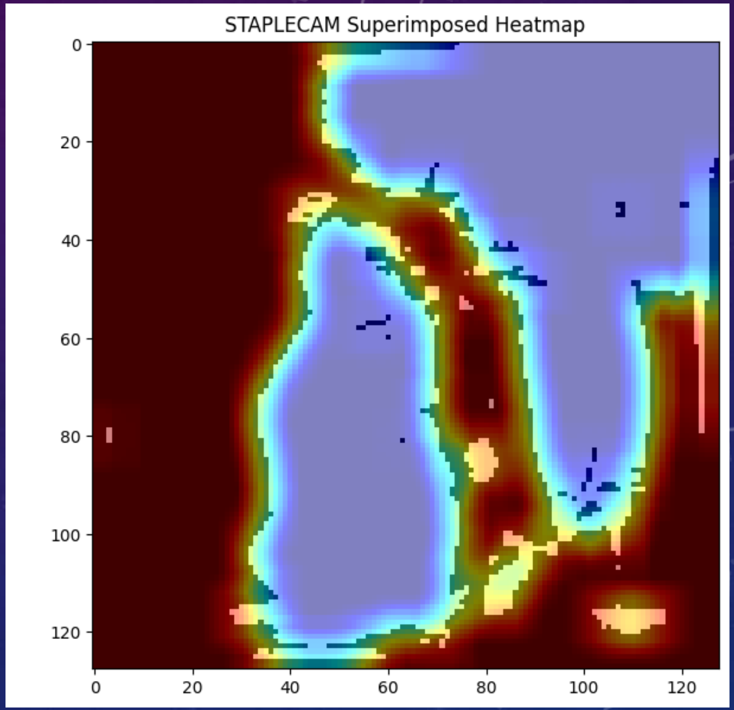


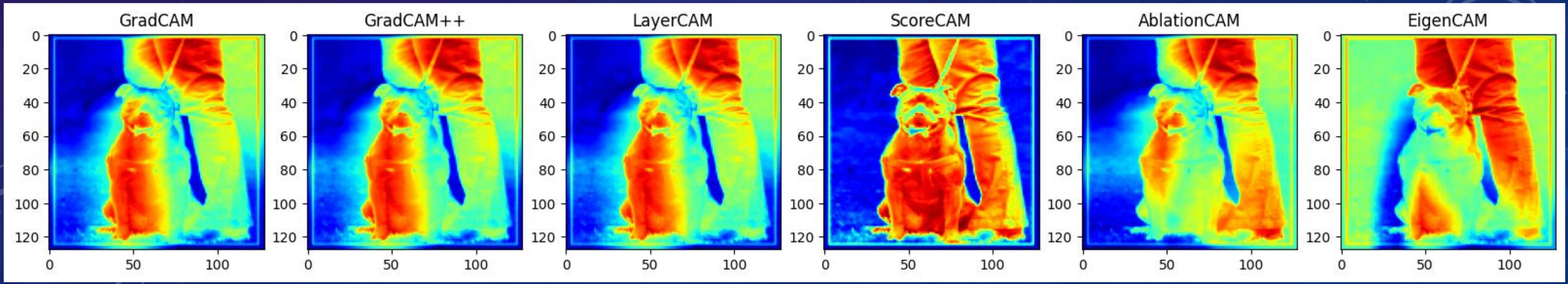
INPUT IMAGE

STAPLE CAM IMAGE

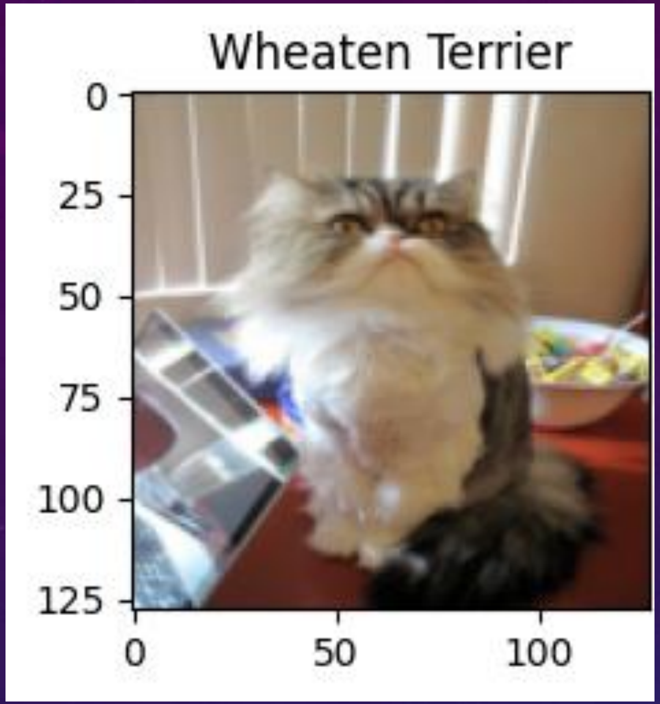


STAPLECAM SUPERIMPOSED HEATMAP

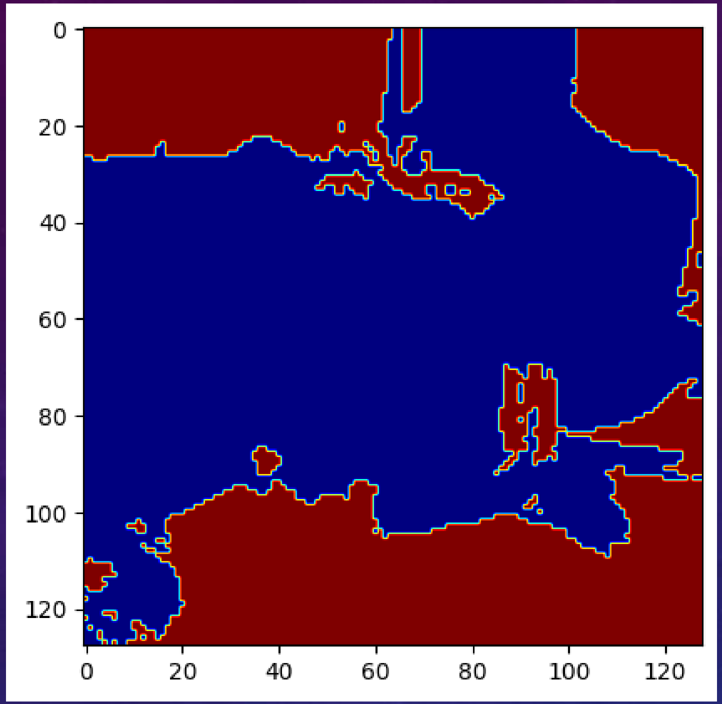




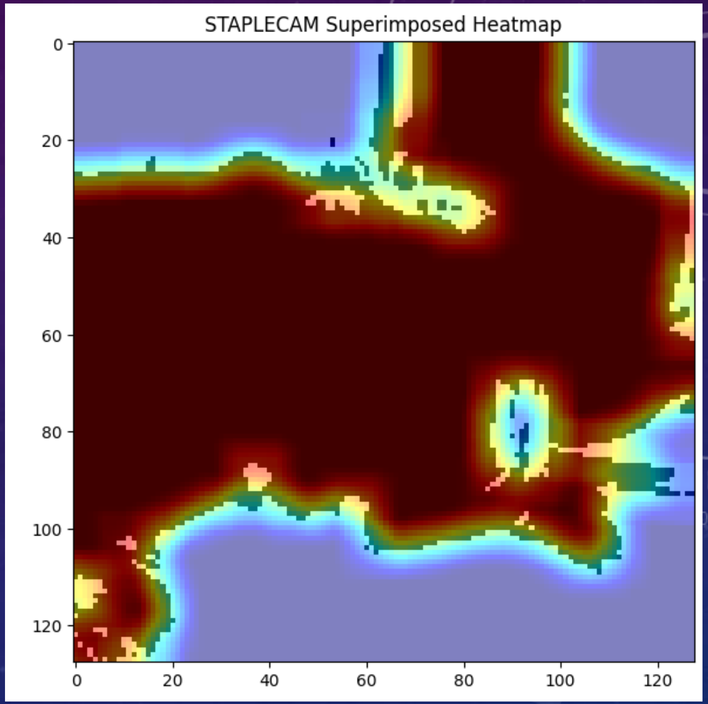
INPUT IMAGE

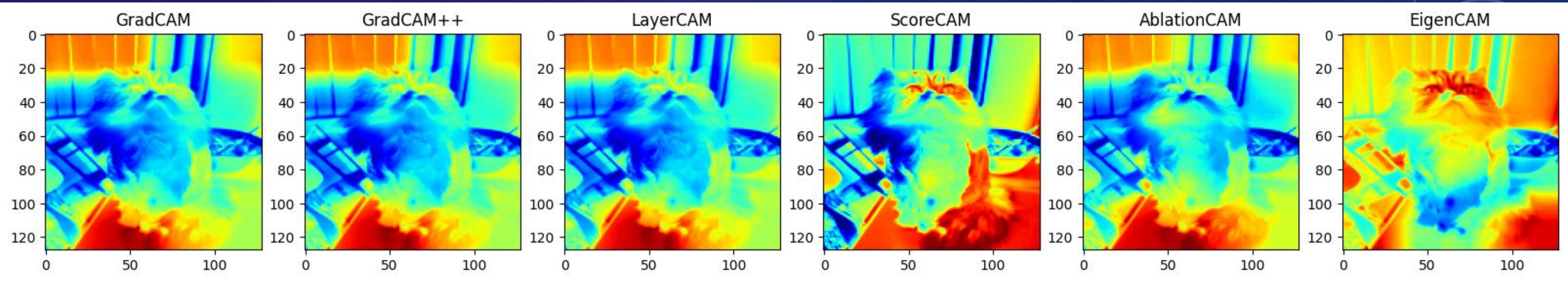


STAPLE CAM IMAGE

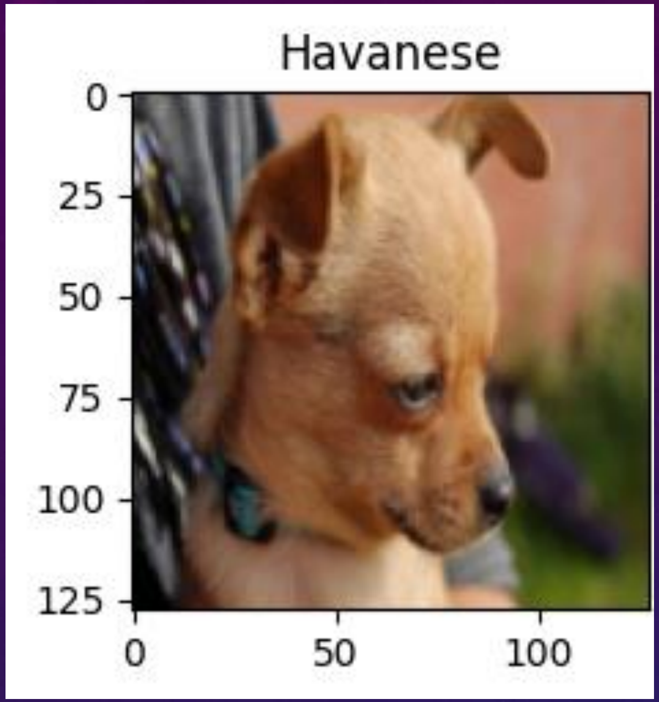


STAPLECAM SUPERIMPOSED HEATMAP

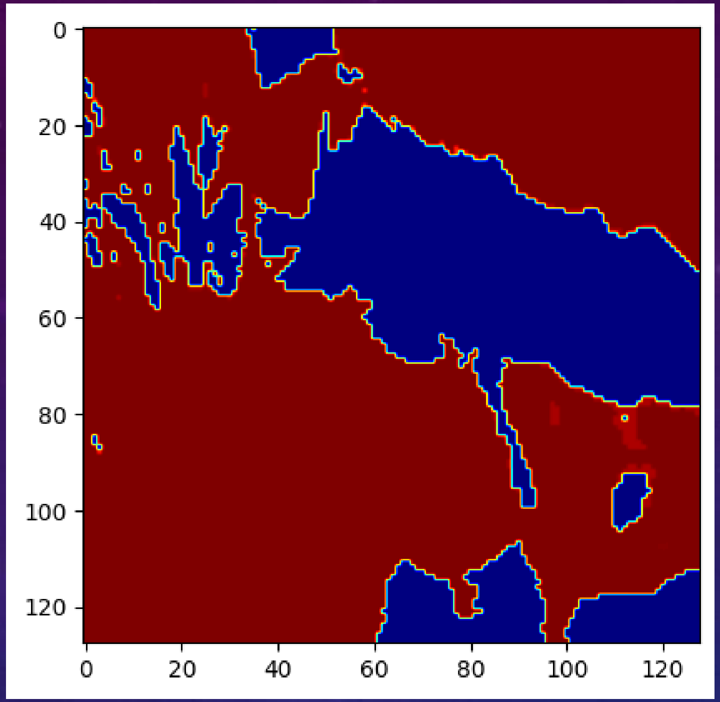




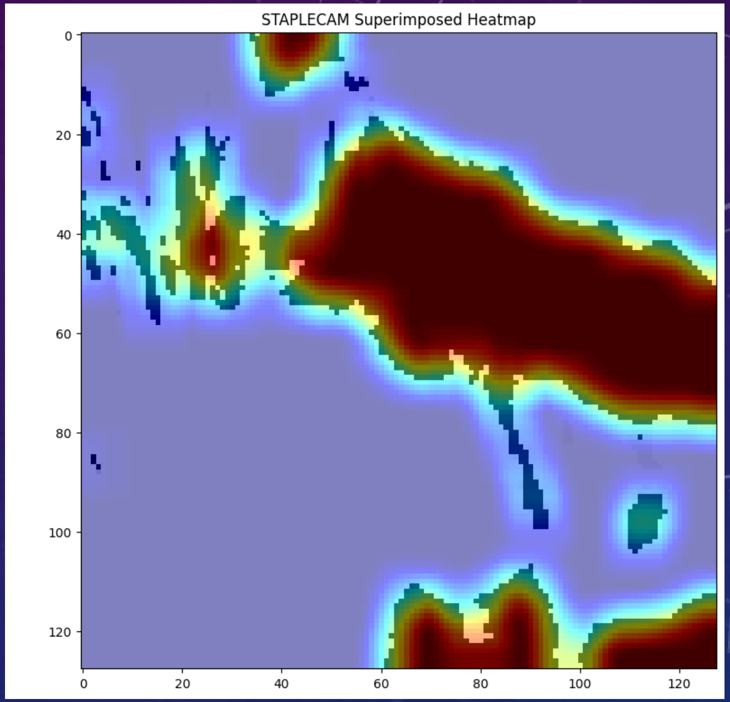
INPUT IMAGE

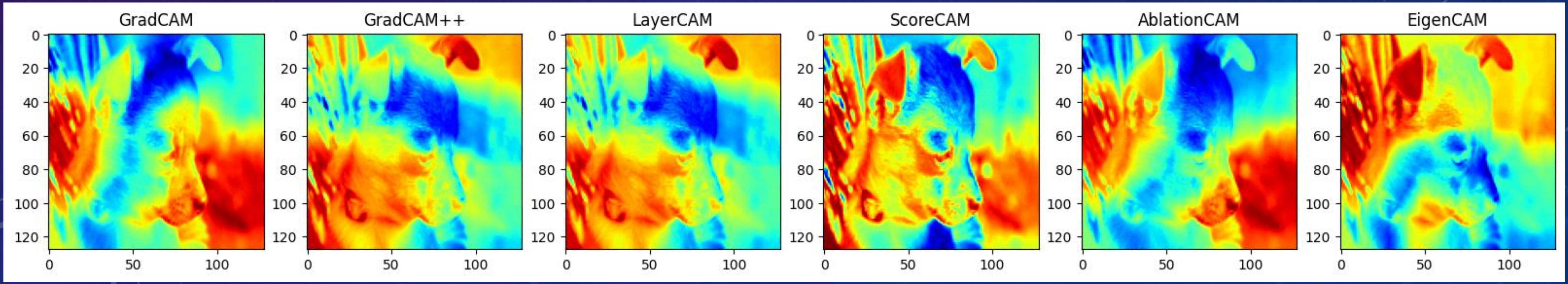


STAPLE CAM IMAGE



STAPLECAM SUPERIMPOSED HEATMAP





THANKS