olution. At each level, our model quadruples the resolution of the image. 2) Our methodology precludes excessive data retrievals and transfers from frequent view changes triggered by a user. 3) Glance is space-efficient and memory-resident at the client side. A single super-resolution model is applicable for any current resolution level with different upscaling rates.

### D. Paper Organization

Section II outlines the background, introducing the structure of geospatial tiled datasets, indexing schemes, and existing image super-resolution methodologies. This is followed by a discussion of our framework and the architecture of its constituent models in section III. Section IV outlines our system architecture and details the interaction and building of each component. Experimental setups, performance benchmarks, and analysis of results are outlined in Section V. Finally, we discuss our conclusions in Section VI.

high-frequency information [19], resulting in blurry images. There are several approaches to improve the perceptual quality of the super-resolved satellite images [19]–[21].

To perform SISR for tiles over multiple resolutions, which facilitate successive interactive zoom operations, the model should support multiple upscaling factors regardless of the zoom level used for the current view. Most existing SISR implementations target only a single upscaling factor. This increases the complexity of model training; each model with different upscaling factors should be trained separately from scratch and results in prolonged training times. Wang et al [22] have proposed incremental training of GANs (by leveraging progressive GANs [7], [23]) that achieve upscaling factors of up to x8. The process trains a model to achieve a magnification level, m, and upon convergence appends an extra set of dense layers to the trained model to train for 2m and reconstructing the image in intermediate steps by progressively performing a





2x upscaling of the input and combining it with the residual output of the model. We harness this incremental curriculum learning approach – leveraging knowledge gained from a lower-level SR network to train for higher SR to simplify the subsequent upscaling inference and speed-up the learning process. Glance also incorporates geospatial metadata into the models to provide conditional information to account for spatial variations in satellite imagery.

## III. Methodology

**Glance** targets space efficiency and memory residency alongside an *in-memory hierarchical tile cache at the client machine* [24] to improve the interactivity of visual analysis. It infers missing high-resolution image tiles locally using a generative deep learning model based on the existing coarse/partial tiles.

### A. Overview of Framework

Glance comprises two key components - 1) **GlanceNet** and 2) **Glance Error Mitigation Module (GEMM)**.

**GlanceNet** generates super-resolution images on-the-fly by using information available in the local cache populated based on historical queries in the form of both low-resolution and partial high-resolution tiles (if any). For a requested image at zoom-level $(n+u)$, i.e., $2^u$ times higher the resolution of an available cached image, all relevant tiles (resolution $(n+u)$) cached as a result of previous browsing, may also be leveraged.

GlanceNet comprises two GAN-based components: the **Super-Resolution** and the **Image Refinement** models as depicted in **Fig. 2**. The Super-Resolution component performs hierarchical geospatial super-resolution based on input images that may be available at different resolutions. The super resolution and image enhancement components work in tandem to generate imagery with high fidelity by enhancing the perceptual quality of super-resolution image. These components are leveraged selectively based on the number of image tiles

high-resolution image (i.e. Partial Super-Resolution Mode) is available in-memory, which is a common phenomenon due to the observed tendency of users to focus on a certain spatiotemporal region during a single visualization sequence [25] leading to persistent areas of interests through OLAP actions such as panning or drill-down.

The main challenge is facilitating multi-level *super-resolution/upscaling* operation for an image tile over diverse *zoom levels*. Diverse topographical characteristics in satellite imagery at different zoom levels adds complexity to the problem. We propose a conditional adversarial network that utilizes *progressive growing of GANs* with additional metadata extracted from $\mathbf{I_{LR}}$ as conditioning information. Finally, the sequential training of all the sub-models (x2,x4 and so on) captures the relations between all levels of cascading resolutions without going through disjoint, prolonged training phases.

**Model Input:** Given a low-resolution RGB image ($I_{LR}$), the input to the super-resolution network is constructed by superimposing the available high-resolution image tiles, if any, with the bicubic interpolation of $I_{LR}$ for the missing regions (this is the super-imposed image $I_{LR+}$), as shown in Fig. 3a.

Every raw satellite image tile-set has additional metadata associated with their domain. We extract the following tile properties to **condition the inputs to both our generator** ($G_1$) **and discriminator** ($D_1$) – (1) zoom level($z$) of $I_{LR}$, and (2) dominant land cover type (based on the National Land Cover Database (NLCD) codes [26]) in terms of $I_{LR}$ pixels (e.g. urban, forest, barren etc.).

The extracted conditional information $g = (z, nlcd\_code)$ is embedded and then passed through a fully connected linear activation layer to resize the embedding to an extra $4^{th}$ metadata channel that gets appended to the 3-channel RGB image ($I_{LR+}$). This helps condition the GAN's training using specific